DERIVING AND APPLYING IMPROVED UPPER BOUNDS FOR
MULTIVARIATE NORMAL PROBABILITY OUTSIDE OF N-CUBES

BY

DONALD R. HOOVER

TECHNICAL REPORT NO. 407
JULY 19, 1988

Prepared Under Contract
N00014-86-K-0156 (NR-042-267)
For the Office of Naval Research

Herbert Solomon, Project Director

Reproduction in Whole or in Part is Permitted
for any purpose of the United States Government

Approved for public release; distribution unlimited.

QUALITY INSPECTED 2

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

AD-A198193

Codes
Dist
Special
A-1

TABLE OF CONTENTS

I. GLOSSARY OF SYMBOLS USED

P_{ij} = correlation coefficient between variables X_i and X_j

$Pr\{\}$ = probability of an event occurring

Σ = correlation matrix

X = n dimensional random variable $\sim N(0,\Sigma)$

X_i = i-th variable component of X

$A_i = Pr\{X_i \notin (-z,z)\}$

$A = Pr\{\bigcup_{i=1}^{n} (A_i)\}$

α_F = Fixed upper bound for $Pr\{A\}$

α_0 = Upper bound for $Pr\{A\}$ for a given z_F

z_F = Fixed edge length of cube for $A_i = Pr\{X_i \notin (-z_F, z_F)\}$

z_0 = Edge length of cube for $A_i = Pr\{X_i \notin (-z_0, z_0)\}$ determined so that $Pr\{A\} \leq \alpha_F$

T = Tree connecting n disjoint events

T_0 = Best bivariate tree connecting n disjoint events such that $\sum_{i=1}^{n} Pr(A_i) - \sum_{ij \in T_0} Pr(A_i \cap A_j)$ is minimized

$B_{ij} = \{A_i \cup A_j\}$ where A_i is linked to A_j on T_0

T' = A tree linking B_{ij} formed by linking $B_{ij}-B_{i'j'}$ iff $i=i'$ or $i=j'$, or $j=i'$, or $j=j'$

ϕ = Standard univariate normal density

$\phi_2^{(p)}$ = Bivariate normal density of $N\left(\begin{pmatrix}0\\0\end{pmatrix}\begin{pmatrix}1 & p\\ p & 1\end{pmatrix}\right)$

$\phi_3 p_1, p_2, p_3$ = Trivariate normal density of $N\left(\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}1 & p_1 & p_2\\ p_1 & 1 & p_3\\ p_2 & p_3 & 1\end{pmatrix}\right)$

DERIVING AND APPLYING IMPROVED UPPER BOUNDS FOR MULTIVARIATE NORMAL PROBABILITY OUTSIDE OF N-CUBES

II. Introduction

The probability areas outside n-dimensional multivariate normal cubes have long been of interest to statisticians. If X is a random vector with an n-dimensional multivariate normal distribution with μ_i the mean of X_i and σ_i the standard deviation of x_i, $i = 1,\dots,n$, then

(1) Pr X is not in the multidimensional rectangle

$$\begin{bmatrix} \mu_1 - z_F\sigma_1 < X_1 < \mu_1 + z_F\sigma_1 \\ \vdots \\ \mu_i - z_F\sigma_i < X_i < \mu_i + z_F\sigma_i \\ \vdots \\ \mu_n - z_F\sigma_n < X_n < \mu_n + z_F\sigma_n \end{bmatrix}$$

is the same as

(2) Pr Y is not the multidimensional cube

$$\begin{bmatrix} -z_F < Y_1 < z_F \\ \cdots \\ -z_F < Y_i < z_F \\ \vdots \\ -z_F < Y_n < z_F \end{bmatrix} \quad \text{where} \quad Y_i = \frac{X_i - \mu_i}{\sigma_i}$$

i.e. Y is X with each component standardized to $N(0,1)$.

This probability is of interest in three cases:

A. Finding the probability, α_0 that a random observation X is not in the region determined by (1) when $u_1, u_2, \ldots, u_n$ are known.

B. Simultaneous confidence interval construction producing confidence intervals for u_i from X_i of the form $[X_i - z_0\sigma_i < \mu_i < X_i + z_0\sigma_i]$ and being at least $(1-\alpha_F)$ confident that all confidence intervals cover the true parameters. This is done when $u_1, u_2, \ldots, u_n$ are unknown.

C. When simultaneously testing the hypothesis $H_1: U_1 = \mu_1$, $H_2: U_2 = \mu_2 \cdots H_n: U_n = u_n$ against the alternative hypothesis $H_{1a}: H_1$ is false $\cdots H_{na}: H_n$ is false, and it is desired to obtain an upper bound α_F for the probability that at least one H_i is rejected when in fact all H_i are true by using acceptance regions constructed the same way as in case B.

In case A, z_F is usually fixed in advance and it is desired to determine α_0 the smallest upper bound for α. In cases B and C, α_F is usually fixed in advance and it is desired to determine the smallest z_0 which will give α_F as an upper bound for the true α.

Current methods used to find α_0 and z_0 in the situations described above are of three types (i) the Scheffé method which incorporates correlations of marginal distributions along the axis of the cubes but builds the confidence intervals for all linear combinations of the marginal axis variables, and this gives extremely conservative (large) values, (ii) the Tukey range and Bonferroni inequality methods which generally give less conservative results but do not incorporate

correlations of marginal distributions along the axis of the cubes.

(iii) A method suggested by Hahn uses the (not necessarily correct) assumption that if $P_{ij} \geq 0$ for all i,j then

$$\int_{-z}^{z} \cdots \int_{-z}^{z} \Phi_n(\Sigma) dx_1 \dots dx_n \geq \int_{-z}^{z} \cdots \int_{-z}^{z} \Phi_n(P) dx_1 \dots dx_n ,$$

where (a) Σ is the covariance matrix of $x_1, \dots, x_n$. WLOG let the variance of all x_i be 1, the diagonal elements of Σ are 1. (b) $P = |P^*|$ where P^* is the minimum over $i \neq j$ of $|P_{ij}|$ the absolute correlation between x_i and x_j. (c) $\Phi_n(P)$ is an n dimensional multivariate normal distribution with $\mathrm{var}(X_i) = 1$ for all i and $\mathrm{cov}(x_i x_j) = P$ for all i,j. A problem with this method is that often when n is large, P^* is likely to be very close to zero giving this method little advantage over a method assuming independence.

The method described in this report is an improved version of the Bonferoni method which can incorporate correlations of the marginal distributions to produce even less conservative estimates for z_0 and α_0 than does the standard Bonferroni method, or any of the other methods mentioned above.

III. Description of the Improved Bonferroni Method

a. Three Pathway Overlap Reduction Theorem. Hunter (1976), Worsley (1982). Represent events $A_1, \dots, A_n$ as vertices $V_1, \dots, V_n$ of a tree whose vertices V_i and V_j are connected are joined by edges e_{ij}, then

$$\text{(3)} \qquad \Pr\left(\bigcup_{i=1}^{n} A_i \right) \leq \sum_{i=1}^{n} \Pr(A_i) - \sum_{i,j: e_{ij} \varepsilon T} \Pr(A_i \cap A_j) .$$

Proof: Since T is a tree, it is always possible to find a permutation $P_1,\ldots,P_n$ of $1,\ldots,n$ so that A_{p_i} is joined to A_{p_j} for some $j < i$ $(i=2,\ldots,n)$. We can then write

$$(4) \qquad \bigcup_{i=1}^{n} A_{p_i} = A_{p_1} \cup (A_{p_2} \sim A_{p_2} \cap A_{p_1}) \cup \cdots \cup$$

$$(A_{p_n} \sim A_{p_n} \cap (A_{p_{n-1}} \cup \cdots \cup A_{p_1})) .$$

The events in (4) are disjoint, so

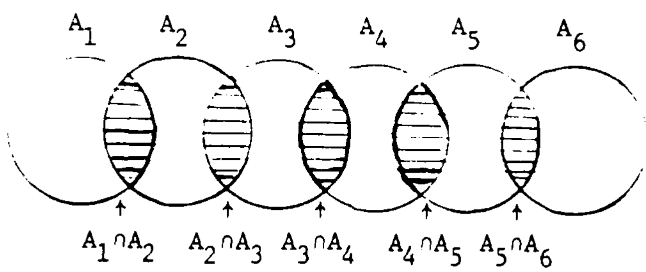
$$(5) \qquad \Pr\left(\bigcup_{i=1}^{n} A_{p_1}\right) = \Pr(A_{p_1}) + \Pr(A_{p_2} \sim A_{p_2} \cap A_{p_1}) + \cdots +$$

$$\Pr(A_{p_n} \sim (A_{p_n} \cap (A_{p_{n-1}} \cup \cdots \cup A_{p_1}))$$

but $\Pr(A_{p_i} \sim A_{p_i} \cap (A_{p_i} \cup \cdots \cup A_{p_{i-1}})) \le \Pr(A_{p_i} \sim A_{p_i} \cap A_{p_j})$ where $1 \le j \le i-1$ hence

$(5) \le \sum \Pr(A_{p_i}) - \sum_{i,j:e_{ij} \in T} \Pr(A_{p_i} \cap A_{p_j})$.

Venn Diagram Explaining Theorem



[Note (A_1 and A_3) (A_2 and A_4), etc. are not necessarily disjoint but are drawn that may be due to inability to represent 6 dimensional intersections on a 2 dimensional diagram.]

This is a tree since there is no loop, i.e.

$$\begin{array}{ccc} A & \rightarrow & B \\ \uparrow & & \downarrow \\ D & \leftarrow & C \end{array}$$

So $\Pr(\bigcup_{i=1}^{6} A_i) \leq \Pr(A_1)+\Pr(A_2)+\Pr(A_3)+\Pr(A_4)+\Pr(A_5)+\Pr(A_6)-\Pr(A_1 \cap A_2)-$
$\Pr(A_2 \cap A_3) - \Pr(A_3 \cap A_4)-\Pr(A_4 \cap A_5)-\Pr(A_4 \cap A_6)$.

A theorem will be presented in a future technical report of this series written by the same author (Hoover 1986) which will enable us, in some cases, to get a lower upper bound for $P_r(\bigcup_{i=1}^{n} A_i)$ than did the previous theorem from incorporation of probabilities of pairwise intersections of events. The improvement (when there is one) will often be very miniscule, and in practice this theorem will be computationally much more difficult to apply than is the three pathway overlap reduction theorem.

IV. Application of Improved Bonferoni Method to Multivariate Normal Distributions.

a. Bivariate Method - Incorporating Only Bivariate and Univariate Normal Distributions.

If $[Y]_n$ is an n-dimensional normalized random vector with mean $[0]_n$ and covariance

$$\begin{bmatrix} 1 & p_{12} \ldots & \ldots & p_{1n} \\ p_{12} & & & \\ \vdots & \ddots & 1 & \ddots \\ p_{1n} & \ldots & \ldots & 1 \end{bmatrix}$$

and $z > 0$. Then let the event

(6.5) $$A = \left\{\begin{matrix} Y_1 - z, Y_1 + z \\ \vdots \\ Y_n - z, Y_n + z \end{matrix}\right\} \text{ does not include } \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

and the event

(6.6) $$A_i = (Y_i - z, Y_i + z) \text{ does not indlude } [0] \; .$$

LEMMA (I) (Sidak, 1968) $\Pr[A_i \cap A_j] = \Pr \left\{\begin{matrix} y_i - z, y_i + z \\ y_j - z, y_j + z \end{matrix}\right\}$ does not include $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a monotonically increasing function of $|p_{ij}|$: that is,

(7) $$|p_{ij}| > |p_{k\ell}| \Leftrightarrow \Pr[A_i \cap A_j] > \Pr[A_k \cap A_\ell] .$$

We should also note that

(7.5) $$\Pr(A_i) = 1 - \int_{-z}^{z} \phi(x)dx \quad \text{for all} \quad i = 1, \ldots, n$$

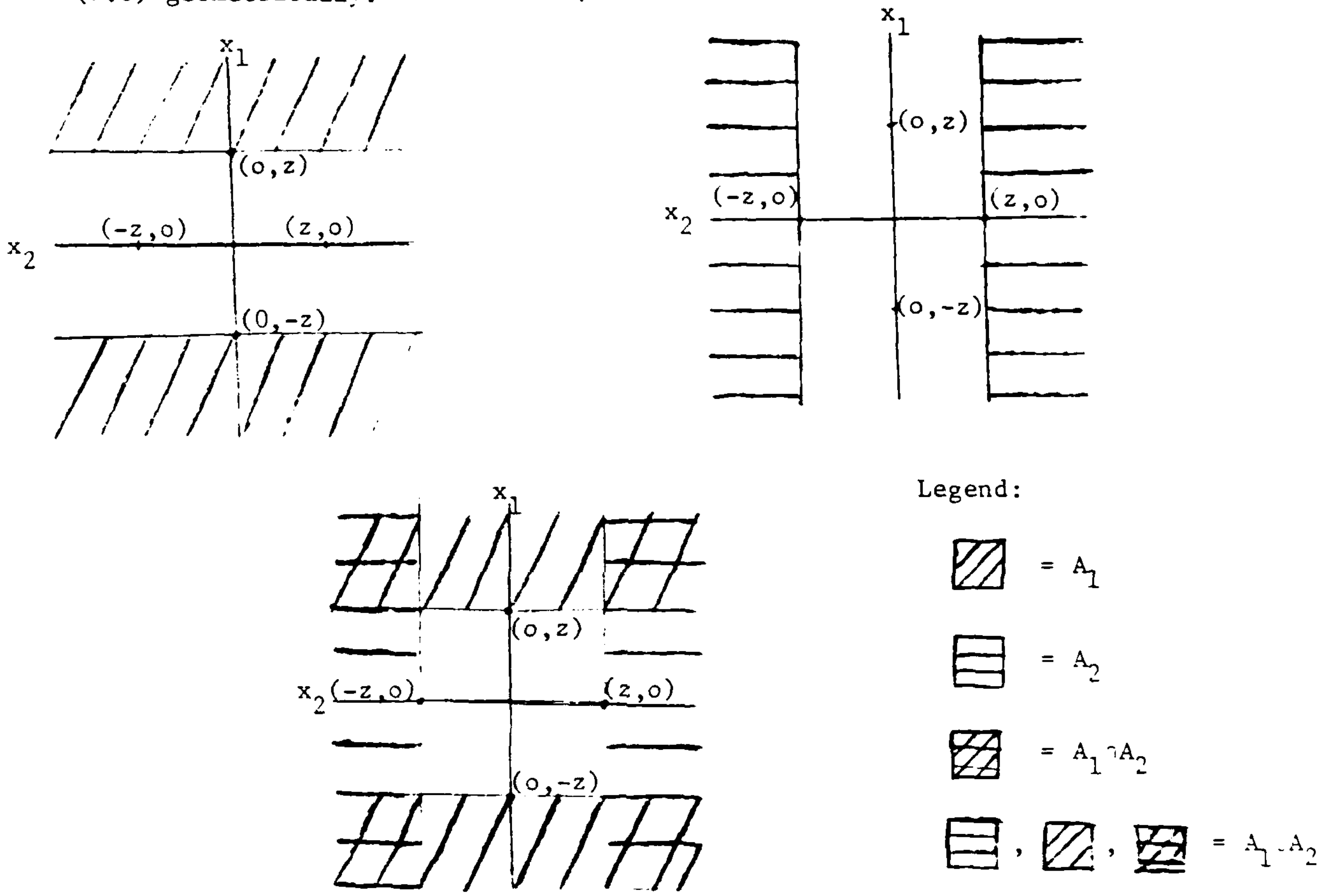
where ϕ is the standard normal density and $\int_{-z}^{z} \phi(x)dx$ can be computed using the IMSL subroutine MDNOR; also

(7.5') $$\Pr(A_i \cap A_j) = \Pr(A_i) + \Pr(A_j) - \Pr(A_i \cup A_j)$$

where $\Pr(A_i)$ and $\Pr(A_j)$ are described in (7.5) above and $\Pr(A_i \cup A_j)$ can be found using the bivariate normal distribution

(7.6) $$\Pr(A_i \cup A_j) = 1 - \int_{-z}^{z} \int_{-z}^{z} \phi_2^{(p)}(x_1, x_2) dx_1 dx_2$$

where $\phi_2^{(\rho)}$ is the bivariate normal density $N\left(\binom{0}{0},\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}\right)$. The integral $\int_{-z}^{z}\int_{-z}^{z}\phi_2^{(\rho)}(x)dx$ can be computed using the IMSL subroutine MDBNOR. The diagrams below explain formulas (7.5) and (7.6) geometrically.



1. Construction of the Best Bivariate Tree

Given any set of multivariate normal variables with a given correlation structure, then using Lemma (I) above and the following theorem below it is possible to construct a "best bivariate tree" over which it is possible to see a "maximal improvement" in upper bounds for all values of z.

Theorem. A best bivariate tree T_0 can be constructed such that if N_{T_0} = a vector of the order statistics of $|P_{ij}|$ over the i and j linked on T_0 then for all $k = (1,\ldots,n)$ and all other trees $\hat{T}$ it will be true that $N_{T_0}(k) \geq N_{\hat{T}}(k)$. This best bivariate tree T_0 is constructed by the following procedure:

(8) Perform the following (n-1) times. Link together any two variables x_i and x_j such that $|P_{ij}|$ is a maximum among all links which have not been established in a previous step and whose establishment will not form a loop with the other links.

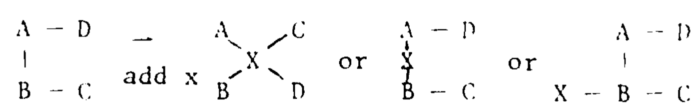
Proof.

FIRST. Given n points, at most $(n-1)$ links can be made between them without forming a circuit, and it is possible to have (n-1) links in a tree no matter how the tree started.

Proof. Induction

(i) Let $n = 2$. Obviously only one link is possible.

(ii) Now if a tree has been established for m items and another item is added. This item can be added anywhere in the tree either by directly connecting to another element of the tree thus adding one link, or by connecting the new element to P elements on the tree and breaking exactly one link in each of the $(P-1)$ segments which connect those P elements.



The (P-1) segments exist since the P elements are connected but it is not P or more segments since there are no circuits. Thus the net number of new links added to the tree by bringing in the new point is P-(P-1) = 1.

SECOND

Now that at most and exactly (n-1) links are possible in every tree it remains to be established that $N_{T_0}(k) \geq N_{\hat{T}}(k)$ for all k. This proof will be inductive. For i = 1 it is obvious that $N_{T_0}(1) \geq N_{\hat{T}}(1)$. Now assume that for i = 1,2,...,k-1 it has been shown that $N_{T_0}(i) \geq N_{\hat{T}}(i)$. Then assume that for k there is a tree $\hat{T}$ such that $N_{\hat{T}}(k) > N_{T_0}(k)$. Let J-0 be the connection in $\hat{T}$ which generates $N_{\hat{T}}(k)$. Then if we consider the first, second,..., (k-1)th links established by the algorithm of this theorem in forming T_0 we must have exactly one string connecting J and 0, i.e. J -....- L - M -....- 0. Otherwise on the kth step the connection J-0 is possible and will be made before a connection generating a lower absolute correlation coefficient is made. k-1 links have been established by the order statistics. N_{T_0} and k-1 different links (some of which may be the same) have been established by the order statistics $N_{\hat{T}}$. The (k-1) links established by the order statistics N_{T_0} must prevent the establishment (excluding redundancy) of any of the k-1 links established by the order statistics $N_{\hat{T}}$. For if there exists a link η-ξ corresponding to the Pth order statistic of $N_{\hat{T}}$: $P \leq k-1$, then $|P_{x_\xi,x_\eta}| \geq |P_{x_j,x_0}|$ > kth order statistic of N_{T_0} and hence the algorithm on the previous page will establish the

link η-ξ on the kth step before establishing a link with a lower absolute correlation.

LEMMA. (II): (k-1) links which do not form a circuit prevent the establishment of any of a set of k-1 different links (excluding redundancy) which themselves collectively do not form a circuit iff each set of the k-1 links is a collection of maximal spanning trees where each maximal spanning tree in one set connects exactly the same points as one of the maximal spanning trees in the other set.

Proof of Lemma.

$\Leftarrow$ WLOG assume each set contains only one maximal spanning tree and that the maximal spanning tree in each set connects exactly the same k points. Then to add any of the links in L_2 which are not in L_1 to the structure formed by the links of L_1 will create a structure with k links between k points and hence must contain a circuit. If there are more than one maximal spanning trees in L_1 and L_2 each maximal spanning tree in L_1 connecting the same points as a maximal spanning tree in L_2. The proof can be extended by considering all the maximal spanning trees of L_1 and adding in the maximal spanning trees of L_2 one at a time.

$\Rightarrow$ If each tree in the set of maximal spanning trees formed by L_1 does not connect exactly to the same points as a corresponding tree in the set of maximal spanning trees formed by L_2, then find a connection in L_2 which connects two points which are not both in the same maximal tree, formed by L_1. Then adding this connection to L_1 will either (i) connect two separate trees together which doesn't form

a circuit or (ii) connect an unlinked point to a tree which doesn't form a circuit or (iii) connect two unlinked points together which does not form a tree. QED.

But this lemma creates a contradiction with the situation we observe since J cannot be connected to 0 due to the first k-1 links of T_0 and each maximal spanning trees formed by the first (k-1) links of T_0 will contain exactly the same points as a maximal spanning tree formed by the first (k-1) links of $\hat{T}$. Thus if it were possible to connect J and 0 together on the kth link of $N_{\hat{T}}$ then it would also be possible to connect J and 0 together on the kth link of T_0 which would have been done by the algorithm forming T_0 since $N_{\hat{T}}(k) > N_{T_0}(k)$.

<u>Implementation of the Procedure Building T_0 on Computer</u>

This can be done by a program which picks the maximum correlation coefficient that has not been used for a link and adds that link to the graph if doing so does not create a loop, i.e. start with $F=0$ and

0) Do while $F < n$.
1) Pick largest remaining $|P_{ij}|$ $i \neq j$.
2) Link i to i if i and j are not in the same tree. Otherwise set $|P_{ij}| = 0$.
3) Note which tree i (or j) has been added to if a variable has been added to a tree. Otherwise if two trees have been merged note that the trees are now the same.
4) $F = F+1$ go to step 0.

This operation involves $0(n)^3$ steps and for 20 or less variables this should not be too complicated.

2. <u>Obtaining the Improved Bonferroni Upper Bounds for Areas Outside Multivariate Cubes Using Formulas (3) and (8).</u>

Let $[X] \sim N(0,\Sigma)$, where (Σ) is a correlation matrix and let T_0 be the best tree obtaind for X using formula (8). If we let A be as defined in (6.5) and A_i as defined in (6.6) for some fixed $z_f > 0$, then

$$\Pr(A) \leq \sum_{i=1}^{n} \Pr(A_i) - \sum_{ij \varepsilon T_0} \Pr(A_i \cap A_j)$$

$$= n\Pr(A_1) - \sum_{ij \varepsilon T_0} \Pr(A_i \cap A_j)$$

(9) by (7.5), (7.6)

$$= n\left[1 - \int_{-z_f}^{z_f} \phi(x)dx\right] - \sum_{ij\varepsilon T_0}\left[\left[1 - \int_{-z_f}^{z_f} \phi(x)dx\right]\right.$$

$$\left. + \left[1 - \int_{-z_f}^{z_f} \phi(x)\,dx\right] - \left[1 - \int_{-z_f}^{z_f}\int_{-z_f}^{z_f} \phi_2^{P_{ij}}(x_1,x_2)dx_1dx_2\right]\right]$$

$$= (2-n)\int_{-z_f}^{z_f} \phi(x)dx + \sum_{ij\varepsilon T_0}\left[1 - \int_{-z_f}^{z_f}\int_{-z_f}^{z_f} \phi_2^{P_{ij}}(x_1,x_2)dx_1dx_2\right] .$$

If α is fixed in advance, say $\alpha = \alpha_F$ and one wishes to determine z_0 such that

$$\Pr(A|z_0) \leq \alpha_F .$$

This can be done iteratively and quite inexpensively by using the Newton-Raphson formula. Start with $(\hat{z}_0)$ an educated guess for z_0. Then let

(9.5) $$\hat{\hat{z}}_0 = \hat{z}_0 - \frac{f(\hat{z}_0)}{f'(\hat{z}_0)}$$

where $f(\hat{z}_0)$ is $[(9)-\alpha_F]$ and

$$f'(\hat{z}_0) = \frac{d}{dz}\Big|_{\hat{z}_0}\left[((9)-\alpha_F)\right] =$$

(*) $$\frac{d}{dz}\Big|_{\hat{z}_0}\left[(2-n)\left[1-\int_{-z}^{z}\phi(x)dx\right] + \sum_{ij\varepsilon T_0}\left[1-\int_{-z}^{z}\int_{-z}^{z}\phi_2^{P_{ij}}(x_1,x_2)dx_1dx_2\right]\right]$$

but

$$\frac{d}{dz}\Big|_{z_0}\left[1-\int_{-z}^{z}\phi(x)dx\right] = -2\phi(\hat{z}_0)$$

and

$$\frac{d}{dz}\Big|_{\hat{z}_0}\left[1-\int_{-z}^{z}\int_{-z}^{z}\phi_2^{P_{ij}}(x_1,x_2)dx_1dx_2\right] = -4\phi(\hat{z}_0)\int_{-\hat{z}_0}^{\hat{z}_0}\phi(x-P_{ij}\hat{z}_0)dx \ .$$

Therefore

(10) $$(*) = (2n-4)\phi(\hat{z}_0) - 4\phi(\hat{z}_0)\sum_{ij\varepsilon T_0}\left[\int_{-\hat{z}_0}^{\hat{z}_0}\phi(x-P_{ij}\hat{z}_0)dx\right] \ .$$

Then set $(\hat{z}_0)$ to equal $\hat{\hat{z}}_0$ and use the above procedure again to produce a better estimate for z_0. It turns out in practice that the Newton-Raphson method $(\hat{z}_0)$ estimate converges quite rapidly to z_0 and is computationally very inexpensive. Note that f is monotonic so the solution is unique.

b. Trivariate Method - Incorporating Trivariate and Bivariate Normal Distributions.

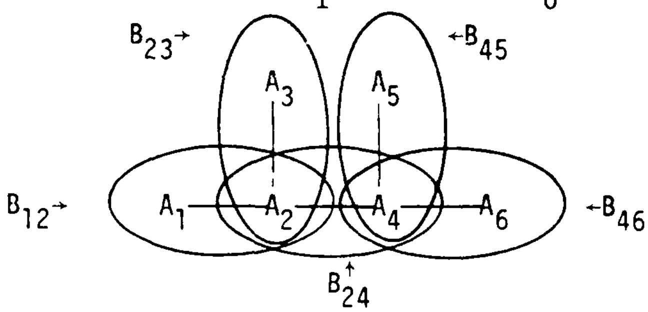
If one is able to integrate over trivariate normal cubes [which can

be done by the IMSL subroutine DMLIN, but is very expensive and time consuming] then one can use the method below to obtain closer upper bounds for z_0 of Statements B, C on page 1 and α_0 of Statement A on page 1 than are obtained by the Bivariate Method.

Let T_0 be the best bivariate tree as constructed in (8). Then for all links $ij \varepsilon T_0$ define the event $B_{ij} = (A_i \cup A_j)$. There will be (n-1) events B_{ij}, since there are (n-1) links on the tree. These events can then be linked together in a new tree T_0' where

(13) $\quad B_{ij}$ can be linked to $B_{i'j'}$ only if i=i' or i=j' or j=i' or j=j' .

For example given events A_i linked by T_0 as below



Then B_{12}-B_{23}-B_{24}-B_{45}-B_{46} is a tree, so is B_{23}-B_{12}-B_{24}-B_{46}-B_{45}.

Except in the case where $A_1 - A_2 - \cdots - A_n$ is a string, there will not be a unique tree linking the B_{ij}'s together such that condition (13) is met.

Assuming that we have a tree of events B_{ij} formed as in the method above with $B_{ij} = (A_i \cup A_j)$ where $X \sim N(0,\ddagger)$ and $\mathrm{Var}(X_i) = 1$ for all i with $A_i = \{|X_i| > z\}$. Then

(14) $$Pr(B_{ij}) = Pr(A_i \cup A_j)$$

given in formula (7.6) and $Pr(B_{ij} \cap B_{jk}) = Pr(B_{ij})+Pr(B_{jk})-Pr(B_{ij} \cup B_{jk})$ where $Pr(B_{ij})$ and $Pr(B_{jk})$ are as defined in (14) and

$$
\begin{aligned}
Pr(B_{ij} \cup B_{jk}) &= Pr(A_i \cup A_j \cup A_k), \\
&= 1 - \int_{-z}^{z} \int_{-z}^{z} \int_{-z}^{z} \phi_3[P_{ij}, P_{jk}, P_{jk}]dx_1 dx_2 dx_3
\end{aligned}
\tag{15}
$$

where $\phi_3[P_{ij}, P_{ik}, P_{jk}]$ is a 3-dimensional normal distribution with

$$
\Sigma = \begin{bmatrix} 1 & P_{ij} & P_{ik} \\ P_{ij} & 1 & P_{jk} \\ P_{ik} & P_{jk} & 1 \end{bmatrix} .
$$

$\phi_3[P_{ij}, P_{ik}, P_{jk}]$ can be claculated with IMSL subroutine DMLIN if Σ is of full rank. Otherwise there is a procedure developed by Uusipaika which can be used.

After incorporating (14) and (15) into (3), it follows that

$$
\begin{aligned}
Pr(\cup A_i) &= Pr\left\{ \bigcup_{e_{ij} \varepsilon T_0'} B_{ij} \right\} \leq \sum_{T_0'} Pr(B_{ij}) - \sum_{ij, i'j' \varepsilon T_0'} Pr(B_{ij} \cap B_{i'j'}) \\
&= \sum_{T_0'} \left[1 - \int_{-z}^{z} \int_{-z}^{z} \phi_2(P_{ij}) dx_1 dx_2 \right] \\
&- \sum_{ij, i'j' \varepsilon T_0'} \left[\left\{ 1 - \int_{-z}^{z} \int_{-z}^{z} \phi_2(P_{ij}) dx_1 dx_2 \right\} \right. \\
&+ \left\{ 1 - \int_{-z}^{z} \int_{-z}^{z} \phi_2(P_{i'j'}) dx_1 dx_2 \right\} \\
&- \left. \left\{ 1 - \int_{-z}^{z} \int_{-z}^{z} \int_{-z}^{z} \phi_3[P_{ij}, P_{ij'}, P_{jj'}] dx_1 x_2 dx_3 \right\} \right]
\end{aligned}
$$

[Assuming WLOG that j=i']

which given the unknown nature of T_0' is the simplest form this can be put in.

Theorem. The trivariate method gives superior results than does the bivariate method, i.e. (16) when applied to a maximal T_0' gives a lower upper bou i for $Pr[A]$ than does (9) applied to T_0.

Proof:

(17) $$\sum_{T_0'} Pr(B_{ij}) = \sum_{i,j\varepsilon T_0} Pr(A_i \cup A_j) = \sum_{i,j\varepsilon T_0} Pr(A_i)+Pr(A_j)-Pr(A_i \cap A_j)$$

$$= \sum_{i=1}^{n} Pr(A_i)(\#\text{ edges connected to } A_i) - \sum_{i,j\varepsilon T_0} Pr(A_i \cap A_j)$$

wlog assume j=i'

(18) $$\sum_{ij,i'j'\varepsilon T_0'} Pr(B_{ij} \cap B_{i'j'}) = \sum_{ij,j'\varepsilon T'_0} Pr(B_{ij} \cap B_{jj'})$$

$$= \sum_{i,j,j'\varepsilon T_0} Pr(A_j)+Pr((A_i \cup A_{j'}) \cap A_j^c)$$

$$= \sum_{i=1}^{n} Pr(A_i)(\#\text{ of edges connected to } A_i)-1$$

$$+ \sum_{i,j,i'\varepsilon T_0} Pr(A_i \cap A_{j'}) \cap A_j^c$$

so subtracting (17) from (18) shows that the upper bound for $Pr(\cup A_i)$ given by (16) applied to T_0' is

$$[\Sigma \Pr(A_i)\cdot[\#\text{ edges connected to } A_i] - \sum_{ij\varepsilon T_0} \Pr(A_i \cap A_j)]$$

$$- [\Sigma \Pr(A_i)[\#\text{ edges connected to } A_i \text{ on } T_0 - 1] + \sum_{jj'\varepsilon T_0'} \Pr(A_i \cap A_{j'}) \cap A_j^c]$$

$$= \Sigma \Pr(A_i) - \sum_{ij\varepsilon T_0} \Pr(A_i \cap A_j) - \sum_{ijj'\varepsilon T_0'} \Pr((A_i \cap A_j') \cap A_j^c)$$

$$\leq \Sigma \Pr(A_i) - \sum_{ij\varepsilon T_0} \Pr(A_i \cap A_j)$$

= the upper bound for $\Pr(\cup A_j)$ given by (0) applied to T_0.

There are some problems with using (16) on T_0' to find an upper bound for $\Pr[\cup A_i]$:

1) It may not be practical except when $A_i,\ldots,A_n$ form a string to find a best tree T_0'. Also the B_{ij} and T_0' formed above do not necessarily give the lowest upper bounds possible for $\Pr(\cup A_j)$ when using trees with bivariate and trivariate normal distributions.

2) It is computationally very expensive and time consuming to calculate the triple interval $\int_{-z}^{z} \int_{-z}^{z} \int_{-z}^{z} \phi(P_{ij},P_{jk},P_{ik})dx_1x_2dx_3$ using the IMSL procedure DMLIN. It may be possible, however, that superior algorithms exist for integrating trivariate normal distributions.

3) When the covariance matrix Σ is not full rank there is no computer algorithm which can calculate the above integral (although Uusipaika's method will work if developed into a computerized procedure).

4) Due to the expensiveness of calculating the trivariate normal intervals it would be impractical to use the Newton-Raphson procedure with this method to find z_0 for a given α_F. Perhaps some other type of numerical estimation could be used, see chapter V part b.

V. Evaluating the Improvements in Upper Bounds for x_0 and Upper Estimates for z_0 Given by Using These Methods.

It is envisioned that these methods could be applied for simultaneous inference on as many as twenty parameters for which we have jointly normal estimates. It has been shown that these methods give closer upper bounds than does the standard Bonferroni method. The standard Bonferroni method is not, however, the best method available for calculating upper bounds for multivariate normal probabilities of lying in cubes centered at the origin. Sidak proved that if x is an $n \times n$ vector distributed $N(\binom{0}{0}, \Sigma)$, where Σ is any $n \times n$ covariance matrix with diagonal elements of 1 then

$$\Pr[\bigcup_{i=1}^{n} x_i \notin (-z,z)] \leq \Pr[\bigcup_{i=1}^{n} y \notin (-z,z)]$$

where y is an $(n \times n)$ vector distributed $N(\binom{0}{0}, I)$. Therefore, by assuming the components of any vector are independent and calculating the probability that the vector lies outside of the n cube centered at the origin, one gets an upper bound for the true probability that the vector lies outside the cube. This upper bound is smaller than the standard Bonferroni upper bound since if we let

$$A_i = \Pr\{x_i \notin (-z,z)\}; \quad A = \bigcup_{i-1}^{n} \{A_i\}$$

$$C_i = \Pr\{y_i \notin (-z,z)\}; \quad C = \bigcup_{i=1}^{n} \{C_i\}$$

Then for all i: $\Pr(A_i) = \Pr(C_i)$ and $\Pr\{C\}$ is the Sidak upper bound for $\Pr(A)$ while $\Sigma \Pr(A_i)$ is the Bonferroni upper bound for $\Pr\{A\}$ and:

$$Pr(A) \underset{Sidak}{\leq} Pr\{C\} \underset{Bonferroni}{\leq} \Sigma Pr(C_i) = \Sigma Pr(A_i) \ .$$

This conservative independence upper bound is easily calculated since $Pr\{C\} = 1-(1-Pr\{C_i\})^n$. The question remains as to when the methods proposed in this report give lower upper bounds than does the conservative assumption of independence.

The answer to this question will be looked at in two different ways. The first will be theoretically to see if we can determine the usefulness of the methods as relates to the $|P_{ij}|i,j\varepsilon T_0$, the values of z_F and α_F and n the number of variables. The second way involves empirical observations of improvement for particular values of n, α_F, z_F and P_{ij} where $P_{ij} = P_{i,j}, ij\varepsilon T_0$.

a. Theoretical Investigation of the Amount of Improvement in Lowering α_0 and z_0 Using Improved Bonferroni Methods.

Theorem 1. For any tree T_0 and any collection $\{P_{ij}\}$ of $P_{ij}, ij\varepsilon T_0$ then given the same tree T_0 and the same $\{P_{ij}\}$ except replacing $P_{i_0j_0}$ with $P'_{i_0j_0}$ where $|P_{i_0j_0}| < |P'_{i_0j_0}|$, the upper bound for $Pr[\cup A_i]$ given using the bivariate and possibly the trivariate method on the new set of $\{P_{ij}\}$ is lower than the upper bound for $Pr[\cup A_j]$ given using the respective method on the original set of $\{P_{ij}\}$.

Proof:

(a) Bivariate: By (Sidak) $(1 - \int_{-z}^{z}\int_{-z}^{z}\phi_2(P_{i_0j_0})dx_1dx_2) > (1 - \int_{-z}^{z}\int_{-z}^{z}\phi_2(P'_{i_0j_0})dx_2dx_2)$ iff $|P_{i_0j_0}| < |P'_{i_0j_0}|$. Thus the upper bound given by (9) will be lower for the new set of $\{P_{ij}\}$.

(b) Trivariate [may not always be true, needs to be proven]. Also need to consider P_{ij}; where $A_i - A_j - A_j$; on the tree.

<u>Theorem 2</u>. For $|P_{ij}| \le P_0$ where $P_0 < 1$, $ij \varepsilon T_0$. Then if we let

A) $\alpha_0(I)$ be the α calculated for a fixed z_F under assumption of independence.

$\alpha_0(B)$ be the α calculated for a fixed z_F using the bivariate method.

$\alpha_0(Tr)$ be the α calculated for a fixed z_F using the trivariate method.

Then let $n \to \infty$. There exists a N_0 such that for all $n > N_0$: $\alpha_F(I) \le \alpha_F(B)$ and $\alpha_F(I) \le \alpha_F(Tr)$.

Also, if we let

B) $z_0(I)$ be the z calculated for a fixed α_F under assumption of independence.

$z_0(B)$ be the z calculated for a fixed α_F using the bivariate method.

$z_0(Tr)$ be the z calculated for a fixed α_F using the trivariate method.

Then let $n = \#$ variables $\to \infty$. There exists a N_0 such for all $n > N_0$: $z_F(I) < z_F(B)$, $z_F(I) < z_F(Tr)$.

<u>Proof of A)</u>.

For fixed z_F the proof is easy since it can be shown that for large enough n, the upper bound given by the trivariate method is greater than 1. This implies, since the bivariate method is never superior to the trivariate method, that the bivariate upper bound is also greater than 1.

Formula (19) gives the trivariate upper bound as

$$\Sigma \Pr(A_i) - \sum_{i,j\epsilon T_0} \Pr(A_i \cap A_j) - \sum_{ij,jk \text{ on } T_0'} \Pr(A_i \cap A_k \cap A_j^c)$$

so adding a new element A_n to the tree adds the terms

$$\Pr(A_n) - \Pr(A_n \cap A_m) - \Pr(A_n \cap A_k \cap A_m^c)$$

$$= [\Pr(A_n) - \Pr(A_n \cap (A_k \cup A_m))]$$

$$= \Pr[(A_n \cap (A_k \cup A_m]^c]$$

and if $|P_{nk}|, |P_{km}|, |P_{nm}| \leq P_0 < 1$ then P_{nk}, P_{km}, P_{nm} ranges over a compact set hence $\Pr[A_n \cap (A_k \cup A_m)^c]$ which is a function of P_{nk}, P_{kn}, P_{nm} and is bounded by zero must achieve its minimum value at some point in this set. But for every point in this set $\Pr(A_n \cap (A_k \cup A_m)^c) > 0$ since $x_n = x_k$ or $x_n = x_m$ is impossible. Therefore $\Pr[A_n \cap (A_k \cup A_m)^c] > \varepsilon > 0$. Thus adding each new element to the tree adds at least ε to the trivariate uppe bound. So if we add $Q > \frac{1}{\varepsilon}$ elements to the tree we will derive the trivariate upper bound above 1 and 1 > (independence upper bound).

Proof of B)

This is very complicated and will not be done.

QED.

Two Part Theorem Regarding the Relationship Between the Improved Bonferroni Lower Bounds for Areas on Normal $(z)^n$ Cubes Centerd at the Origin and the Value of z.

Part I) As $z \downarrow 0$ for a given $n > 3$ and nonsingular correlation sub-matrix $(\Sigma)'_q$ of size $q > 3$, and $x \sim N\left(\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}, (\Sigma)_n\right)$, there exists a z_L (depending on n and Σ) such that $z < z_L$ implies that the upper bounds given for $\Pr[X_i \notin (-z,z)$ any $i]$ are greater than 1 for both the bivariate and the trivariate methods.

Proof:

It has been shown already that the bivariate upper bound can be no lower than the trivariate upper bound so it suffices to prove the statement for the trivariate upper bound only.

Let H and K be two events such that H is not linked to K on T_0; a link $B_{hj}-B_{jk}$ does not exist on T_0' for any j; and there exist variables X_n, X_0 such that $[X_j, X_k, X_n, X_0]$ has a nonsingular distribution. Then the trivariate upper bound for the $\Pr[X_i \notin (-z,z)$ any $i]$ will equal

$$\Pr\{\bigcup_{i=1}^{n} X_i \notin (-z,z) i=1,\ldots,n\}$$

$$+ \Pr\{\bigcap_{i\neq h,k} (X_i \in (-z,z)) \cap X_h \notin (-z,z) \cap X_k \notin (-z,z)\}$$

+ Probabilities of other, disjoint terms

$$\geq [\Pr\{\bigcap_{i\neq h,k} (X_i \in (-z,z)) \cap X_h \notin (-z,z) \cap X_k \notin (-z,z)\}] +$$

$$[1 - \Pr\{\bigcap_{i=1}^{n} X_i \in (-z,z)\}]$$

$$= 1 + [\Pr\{\bigcap_{i\neq h,k} (X_i \in (-z,z)) \cap X_h \notin (-z,z) \cap X_k \notin (-z,z)\}$$

$$- \Pr\{\bigcap_{i=1}^{n} X_i \in (-z,z)\}]$$

But as $z \downarrow 0$ the ratio

$$\frac{\left[\Pr\{\bigcap_{i\neq h,k} (X_i \varepsilon(-z,z)) \cap X_h \notin(-z,z) \cap X_k \notin(-z,z)\} + \Pr\{\bigcap_{i=1}^{n} X_i \varepsilon(-z,z)\}\right]}{\Pr\{\bigcap_{i=1}^{n} X_i \varepsilon(-z,z)\}}$$

$$= \frac{\Pr\{\bigcap_{i\neq h,k} X_i \varepsilon(-z,z)\}}{\Pr\{\bigcap_{i=1}^{n} X_i \varepsilon(-z,z)\}} \Rightarrow \frac{\text{density on the h,k plane passing through origin}}{\text{density on origin}}$$

$= \infty$ due to nonsingularity of distribution of (X_p, X_0, X_h, X_k).

Therefore there exists a value of z_L such that $z < z_L \Rightarrow$

$$\frac{\Pr\{\bigcap_{i\neq h,k} X_i \varepsilon(-z,z)\}}{\Pr\{\bigcap_{i=1}^{n} X_i \varepsilon(-z,z)\}} > 2$$

$$\Rightarrow 1 + \Pr\{\bigcap_{i\neq h,k} (X_i \varepsilon(-z,z)) \cap X_h \notin(-z,z) \cap X_k \notin(-z,z)\} - \Pr\{\bigcap_{i=1}^{n} X_i \varepsilon(-z,z)\} > 1$$

$\Rightarrow$ trivariate upper bound for $\Pr\{z_i \notin(-z,z) \text{ any } i\} > 1$

QED.

Part II) As $z \uparrow \infty$. For a given n and nondegenerate (i.e., no two events are the same) correlation matrix Σ and $X \sim N(\binom{0}{0}, [\Sigma]_n)$ there exists an z_μ depending on n and Σ such that $z > z_\mu$ implies that if

Q_1 is the upper bound estimate of the probability (X) is not in the cube centered at zero with edge length z using independence.

Q_2 is the upper bound estimate of the probability (X) is not in the cube centered at zero with edge length z using either the bivariate or the trivariate method.

Then

$$\frac{Q_1}{Q_2} < (1+\varepsilon) .$$

Proof. [For bivariate method].

1) First we will show that for a given tree of size n

$$\lim z \uparrow \infty \left[\sup_{i \neq j} \frac{\Pr(A_i \cap A_j)}{\Pr(A_i)} < \frac{\varepsilon}{2n}\right] \text{ for all } \quad \varepsilon > 0$$

where

$$A_i = \Pr\{X_i \notin (-z,z)\}$$

$$A_j = \Pr\{X_j \notin (-z,z)\}.$$

One way to do this is to note that if

$$\begin{pmatrix} X_i \\ X_j \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}\left(\begin{pmatrix} 1 & P_{ij} \\ P_{ij} & 1 \end{pmatrix}\right)\right)$$ Remember $|P_{ij}| \neq 1$ since Σ is nondegenerate.

WLOG assume $P_{ij} \geq 0$.

Then one can say $X_j = P_{ij}X_i + \sqrt{1-P_{ij}^2}(y)$ where y is $N(0,1)$ and independent of X_i. Now for all $\varepsilon > 0$ and $\sigma > 0$, $\exists z_b$ such that for any $z > z_b$

$$(*) \qquad \Pr\{|X_i| > z+\sigma \,|\, |X_i| > z\} < \left(\frac{\varepsilon}{8n}\right) .$$

So for any $z > z_b$

$$\Pr\{|X_j| > z \,|\, |X_i| > z\}$$

$$\leq \Pr\{|X_i| > z+\sigma \,|\, |X_i| > z\} + \Pr\{|X_j| \varepsilon [z, z+\sigma] \,|\, |X_i| > z\}$$

$$\leq \Pr\{|X_i| > z+\sigma \,|\, |X_i| > z\} + \Pr\{|X_j| \varepsilon [z, z+\sigma] \,|\, |X_i| > z\}$$

$$\leq \frac{\varepsilon}{8n} + \Pr\left\{y < \frac{(-1-P_{ij})}{\sqrt{1-P_{ij}^2}}\, z\right\} + \Pr\left\{y > \frac{z-P_{ij}(z+\sigma)}{\sqrt{1-P_{ij}^2}}\right\}$$

$$< \frac{\varepsilon}{8n} + \Pr\left\{y < -\left(\frac{1+P_{ij}}{\sqrt{1-P_{ij}^2}}\right) z\right\} + \Pr\left\{y > \frac{(1-P_{ij})}{\sqrt{1-P_{ij}^2}}\, z - \frac{\sigma P_{ij}}{\sqrt{1-P_{ij}^2}}\right\}$$

but $\exists z_c$ such that $z > z_c \Rightarrow$

$$(**) \qquad \Pr\left\{y < \frac{-(1+P_{ij})}{\sqrt{1-P_{ij}^2}}\, z\right\} < \frac{\varepsilon}{8n}$$

and

$$(***) \qquad \Pr\left\{y > \frac{1-P_{ij}}{\sqrt{1-P_{ij}^2}}\, z - \frac{\sigma P_{ij}}{\sqrt{1-P_{ij}^2}}\right\} < \frac{\varepsilon}{8n}$$

So if $z > \max(z_b, z_c)$ then (*), (**) and (***)

$$\Rightarrow \Pr\{|X_j| > z \,|\, |X_i| > z\} \leq \frac{3\varepsilon}{8n} \leq \frac{\varepsilon}{2n}$$

$$\Rightarrow \Pr\{|X_j| > z \cap |X_i| > z\} < \frac{\varepsilon}{2n} \Pr\{|X_j| > z\}$$

II) So WLOG let $P_{ij} = \sup_{k \neq \ell} |P_{k\ell}|$ for all $k,\ell \varepsilon T_0$ and z be larger than $\max(z_b, z_c)$. Then

$$\Pr\{|X_k| > z \cap |X_\ell| \geq z\} < \frac{\varepsilon}{2n} \Pr\{|X_j| > z\}$$

$$\Rightarrow Q_2 = \sum_{k=1}^{n} \Pr\{|X_k| \geq z\} - \sum_{k,\ell \varepsilon T_0} \Pr\{|X_k| > z \cap |X_\ell| > z\}$$

$$\geq (1 - \frac{\varepsilon}{2n}) \sum_{T_0} \Pr\{|X_k| > z\}$$

Now

$$\left[\sum_{T_0} \Pr\{|X_k| > z\}\right] \geq Q_1 \underset{\text{independence}}{\Pr} \{\bigcup_{k=1}^{n} |X_k| > z\}$$

$$\Rightarrow \frac{Q_1}{Q_2} < \frac{\Sigma_{T_0} \{\Pr|X_k| > z\}}{Q_2} < \left[\frac{1}{1 - \frac{\varepsilon}{2n}}\right] < (1+\varepsilon) \text{ for } \varepsilon < 1 .$$

The proof for the trivariate method follows from the proof of the bivariate method, since:

(I) We saw in the bivariate method

$$\lim_{z \uparrow \infty} \left[\sup_{i \neq j} \frac{\Pr(A_i \cap A_j)}{\Pr(A_i)} \leq \frac{\varepsilon}{2n}\right] \text{ for all } \varepsilon < 0 ,$$

which implies that

$$\lim_{z\uparrow\infty}\left[\sup_{i\neq j\neq j'}\frac{\Pr(A_i\cap A_j^c\cap A_j')}{\Pr(A_j)}\right]\leq\frac{\varepsilon}{2n}\text{ for all }\varepsilon>0\,.$$

(II) So the trivariate upper bound (Q_2) given by (19) is

$$(*)\qquad \Sigma\Pr(A_i)-\sum_{ij\varepsilon T_0}\Pr(A_i\cap A_j)-\sum_{ijj'\varepsilon T_0'}\Pr((A_i\cap A_{j'})\cap A_j^c)\,.$$

but from the proof of the bivariate method and the fact that $\Pr((A_i\cap A_j')\cap A_j^c)\leq\Pr(A_i\cap A_j')$ it follows that for z large enough

$$(*)\quad\geq n\Pr(A_i)-(n-1)\frac{\varepsilon}{2n}\Pr(A_i)-(n-2)\frac{\varepsilon}{2n}\Pr(A_i)$$

$$= n\text{PR}(A_i)-\frac{(2n-3)\varepsilon}{2n}\Pr(A_i)\leq(n-\varepsilon)\Pr(A_i)$$

$$=(1-\frac{\varepsilon}{n})\sum_{T_0}\Pr\{|X_k|>z\}\,.$$

(III) Now

$$\sum_{T_0}\{\Pr|X_k|>z\}>Q_1\underset{\text{independence}}{\geq}\Pr(\bigcup_{k=1}^{n}|X_k|>z)$$

$$\rightarrow\frac{Q_1}{Q_2}<\frac{\Sigma_T\{\Pr|X_k|>z\}}{Q_2}<\left[\frac{1}{1-\frac{\varepsilon}{n}}\right]<(1+\varepsilon)\text{ for }n>2\,.$$

QED

Conclusion.

The previous theorems of this and other sections give us

(a) The trivariate method always gives lower upper bounds than does the bivariate method.

(b) Everything else remaining the same (i.e. n, T_0,Σ) replacing a $P^{(1)}_{i_0j_0}$ on the tree with a $[P^{(2)}_{i_0j_0}]$ such that $|P^{(1)}_{i_1j_0}| < |P^{(2)}_{i_0j_0}|$ will always produce a lower upper bound using the bivariate method and will probably also produce a lower upper bound using the trivariate method.

(c) For fixed n, nonsignular Σ, T_0 and T' making z_0 close to zero causes the bivariate and trivariate methods to produce higher upper bounds in relation to those produced by the assumption of independence.

(d) For fixed n, nonsingular Σ, T_0 and T' making z_0 very large causes the bivariate and trivariate methods to produce upper bounds which get infinitely closer to those produced by the conservative assumption of independence.

(e) For bounded $|P_{ij}|$ $\forall_{i,j}$; increasing n eventually causes the bivariate and trivariate methods to produce poorer upper bounds with respect to those produced by the conservative assumption of independence.

Thus we can see there is a complex relation involving n, (z_F or α_F), and the P_{ij} $1 \leq i,j \leq n$ which governs the effectiveness of using the trivariate/bivariate methods over the assumption of independence. In the next section we try to develop an actual picture of this effectiveness using sample calculations of upper bounds for α_0 given z_F and z_0 given α_F for selected values of n, and $P_{ij} i,j \varepsilon T_0$.

b. Actual Upper Bounds Obtained for $z_0|\alpha_F$ and $\alpha_0|z_F$ From Using the Trivariate and Bivariate Methods.

1. Table 1 contains upper and exact bounds for $\alpha_0|z_F$ obtained by using several methods involving several assumptions on z_F, n, and Σ:

(i) z_F was taken to be 1.96, 2.25, 2.5 and felt that these values of z were representative of what would be of interest in practice.

(ii) n was taken to be 4,5,6,...20 because it was felt that in practice, simultaneous inference on 4 to 20 variables would be desired.

(iii) The methods used were

A. Column 3: Conservative assumption of independence [Corollary 3, Sidak, 1971] which says for $X \sim N(0,\Sigma)$

$$\Pr\{\bigcup_{i=1}^{n} X_i \notin (-z_F\sigma_i, z_F\sigma_i)\} \leq 1 - \prod_{i=1}^{n} \Pr\{X_i \varepsilon (-z_F\sigma_i, z_F\sigma_i)\}$$

B. Column 4: Bivariate method assuming the $|P_{x_ix_j}|$ all have the same value P for $e_{ij}\varepsilon T$. The value P was allowed to be .1,.2,...,.9,.99. Calculations were done by a computer program using formula (9) and the IMSL procedures MDNOR and MDBNOR to calculate bivariate and univariate normal cube probabilities respectively

C. Column 5: Trivariate method making the same assumptions as in B but also assuming $P_{x_ix_{j'}} = (P)$ for $e_{ij,jj'}\varepsilon T'$. Calculations were done by computer programs using formula (16) and the IMSL procedures MDBNOR and DMLIN to calculate trivariate and bivariate cube probabilities respectively.

D. Column 6: The exact value for the equicorrelation case where $P_{k\ell} \equiv P$ for all k,ℓ. The formula taken from Hahn and Henderson is:

$$\Pr\{\bigcup_{i=1}^{n} A_i\} = 1 - \left[\int_{-\infty}^{\frac{z_F+\sqrt{p}\,y}{\sqrt{1-p}}} \phi(x)dx - \int_{-\infty}^{\frac{-z_F+\sqrt{p}\,y}{\sqrt{1-p}}} \phi(x)dx\right]^n \phi(y)dy$$

Again P was allowed to be .1,.2,...,.9,.99

E. Column 7: Trivariate method making the same assumptions as in B, but also assuming $P_{x_i x_{j'}} = (P)^2$ for $e_{ij,jj'} \varepsilon T'$. This is an AR(1) time series process. Calculations were done as in Column 5.

Examination of this table will give some ideas as to the improvement of the upper bound for $\alpha_0 | z_F$ which are obtained through the bivariate and trivariate methods. The two points of interest are:

1. How much lower, if any, are the upper bounds for α_0 obtained from the bivariate and trivariate estimates than are those obtained from the conservative assumption of independence.

2. How close are the bivariate and trivariate upper bound estimates to the actual values.

Column 6 in Table 1 gives actual values of α_0 for the case where there is equicorrelation among all the variables, i.e.

$$\Sigma = \begin{matrix} 1 & & & & \\ P & 1 & & & \\ P & P & 1 & & \\ \vdots & P & \ddots & \ddots & \\ P & \cdots & P & P & 1 \end{matrix}$$

Column 3 gives the independence upper bound for the α_0 in column 6 while column 4 is the bivariate upper bound for the α_0 in column 6. Column 5 is the trivariate upper bound for the α_0 in column 6.

The pattern observed in these tables is far too complex to be described in the limited space of this section. But some general observations should be made.

a. For correlations of less than .3, the bivariate and trivariate methods do not improve much or at all over the conservative assumption of independence for all values of n and z_F examined.

b. For $P_{ij} \geq .6$, $z_F \geq 2.25$ and all n, the bivariate upper bounds for α_0 are much smaller than are the independence upper bounds. Example $z_F = 2.75$, $P_{ij} = .8$, $n = 18$ the independence upper bound for α_0 is .10201, while the bivariate upper bound for α_0 is .07575.

c. For all P_{ij}, $z_F \geq 2.25$ and all n the trivariate upper bound for the AR(1) process is not substantially smaller than is the bivariate upper bound. Example $z_F = 2.25$, $P_{ij} = .4$, $P_{ij'} = .16$ and $n = 6$. The bivariate upper bound for α_0 is .13502 and the trivariate upper bound for α_0 is .13254.

d. For $P_{ij} \geq .4$ and all z_F and n, the trivariate upper bound for α_0 in the equicorrelated case is significantly lower than is the bivariate upper bound for α_0, yet is much larger than the actual value of α_0 especially if P_{ij} is large. Example 1: $P_{ij} = .5$, $n = 8$, $z_F = 2.5$.

The bivariate upper bound is .08998, the trivariate upper bound is .08400 and the actual value is .07456. Example 2: $P_{ij} = .7$, $n = 13$, $z_F = 2.25$. The bivariate upper bound for α_0 is .23582, the trivariate upper bound is .19754 and the actual value is .13471.

For the high correlations, i.e. $P_{ij} \geq .7$ the trivariate and the bivariate methods give better upper bounds for α_0 for n even as large as 20 than does the conservative assumption of independence. In the case where the correlations between variables become multiplicatively smaller the farther away one moved along the tree (i.e. AR(1)) the reductions in the upper bounds from using the trivariate method were not much greater than were those from using the bivariate method. If the correlations of variables far apart from each other on the tree are high with respect to the correlations between adjacent variables on the tree (equi-correlated case) then even the trivariate upper bounds for α_0 were much higher than the true values of α_0.

2. Tables 2A, 2B and 2C contain upper bounds for $z_0{}'\alpha_F$ obtained by using the bivariate method along with several assumptions on α_F, n and Σ.

(i) α_F was taken to be .05, .10, and .20 because it was felt that in practice jointly simultaneous confidence intervals of 80%, 90% and 95% probability would be desired.

(ii) n was taken to be 4,5,6,...,20 because it was felt in practice that simultaneous inference on 4 to 20 variables would be desired.

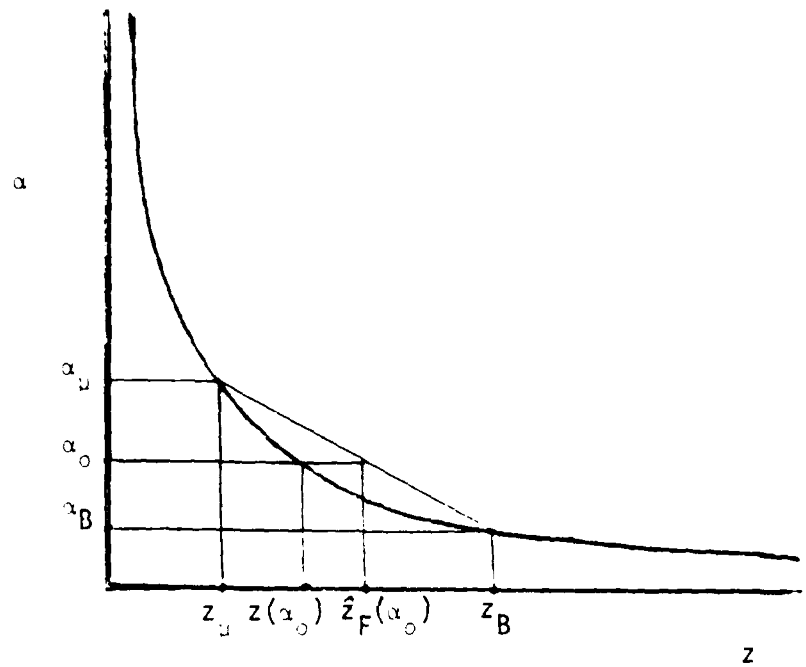
(iii) (Σ) was chosen so that $|P_{X_j X_j}|$ $e_{ij} \varepsilon T$ always equaled some constant P. P was set at .1,.2,...,.9,.95 and .99.

$z_0 | \alpha_F$ was calculated by choosing an initial $\hat{z}_0$ and then calculating $\alpha_{\hat{z}_0}$ by formula (9) using the IMSL procedures MDNOR and MBDNOR. A new $\hat{z}_0$ was calculated by (9.5). This was repeated until $|\alpha_{\hat{z}_0} - \alpha_F| < .0001$ at which point z_0 was set to $\hat{z}_0$.

No attempt was made in this table to find trivariate upper bounds for $z_F | \alpha_0$. Due to the present computational difficulty of working with trivariate intervals, it would be difficult to use the Newton-Raphson procedure to do this as could be done for bivariate upper bounds. Some type of interpolation method would probably be the best way to obtain these bounds. One crude method would be by finding z_B (the bivariate upper bound for z_0. Remember $z_B > z_T$. Calculate (under the trivariate method) $\alpha_B | z_B$. It follows that $\alpha_B < \alpha_0$. Find z_u such that the single interval probability that $X_i \notin (-z_u, z_u)$ is α_0. Then calculate $\alpha_u | z_u$ for all of the intervals together using the trivariate method. It will follow that $\alpha_u > \alpha_0$. Finally let

$$\hat{z}_F = \frac{\alpha_u - \alpha_0}{\alpha_u - \alpha_B} z_B + \frac{\alpha_0 - \alpha_B}{\alpha_u - \alpha_B} Z_u \quad . \tag{19}$$

The $\hat{z}_F$ we obtain will be smaller than that obtained by the bivariate method. It also should be larger than the true trivariate upper bound z_F since it makes sense to believe that $\hat{z}_\alpha$ is a convex function, i.e.



As was the case in Table 1; the patterns shown in Tables 2A, 2B and 2C are too complex to describe completely. Some important things can and should be noted from these tables, however. First for α_F = .05, .10 and .20; the bivariate method generally gives lower values of z_0 for all n than does the conservative assumption of independence. Although the reduction in values may seem small they are actually larger when one considers the reduction in the hyper-volume of the cube as opposed to the reduction in the edge length of the cube.

Example 1: $\alpha_F = .20$, $n = 6$, $P_{ij} = .4$.

The conservative assumption of independence gives an upper bound for z_0 of 2.111 while the bivariate method gives an upper bound for z_0 of 2.087. This representates a reduction of edge length to .9886 of its original length. But it also represents a reduction in the hyper volume of the cube to (.9886)**6 = .9337 of its original value.

Example 2: $\alpha_F = .20$, $n = 10$, $P_{ij} = .7$.

The conservative assumption of independence gives an upper bound for z_0 of 2.309 while the bivariate method gives an upper bound for z_0 of 2.213. This represents a reduction in edge length to (.9584) of its original length. But it represents a reduction in the hyper volume of the cube to $(.9584)^{**10} = .6540$ of its original volume.

Example 3: $\alpha_F = .05$, $n = 15$, $P_{ij} = .9$.

The conservative assumption of independence gives an upper bound for z_0 of 2.932 while the bivariate method gives an upper bound for z_0 of 2.742. This represents a reduction of edge length to (.9352) of its original length and a reduction in hyper volume of the cube to $(.9352)^{**15} = (.3660)$ of its original volume

Example 4: $\alpha_F = .10$, $n = 20$, $P_{ij} = .99$.

The conservative assumption of independence gives an upper bound for z_0 or 2.799 while the bivariate method gives an upper bound for z_0 of 2.216 which is .7917 of its original length and a reduction in hyper cube column to $(.7917)^{**20} = .00936$ of its original volume.

Table 3 compares trivariate and bivariate upper bounds for $z_0 | \alpha_F = .10$ for $n = 5,10$ and 20 and $P_{ij} = .6,.7,.8$ and .9. The bivariate upper bounds are from Table 2B. The trivariate upper bounds are calculated as in (19) using both the case where $P_{ij}{}' = (P_{ij})$ $e_{ij,jj}{}',\varepsilon T'$ and the case where $P_{ij'} = (P_{ij})^2$ $e_{ij,jj}{}',\varepsilon T'$. The trivariate method does give significant reductions below the bivariate method in the case where $P_{ij'} = P_{ij}$.

Example 1: $\alpha_F = .10$, $n = 5$, $P_{ij} = .7$, $P_{jj'} = .7$.

The trivariate upper bound for z_0 is 2.17 while the bivariate upper bound for z_0 is 2.23. This gives an edge length ratio reduction of .973 and a hypercube volume reduction of $(.973)^{**5} = .872$.

Example 2: $\alpha_F = .10$, $n = 20$, $P_{ij} = .9$, $P_{jj'} = .9$.

The bivariate upper bound for z_0 is 2.59 while the trivariate upper bound for z_0 is 2.49. This gives an edge length ratio of .961 and a hyper volume reduction of $(.961)^{**20} = .455$.

In the case where $P_{ij'} = (P_{ij})^2$ the reductions given by the trivariate method are not as impressive.

Example 3: $\alpha_F = .10$, $n = 10$, $P_{ij} = .8$, $P_{jj'} = .64$.

The bivariate upper bound for z_0 is 2.44 while the trivariate upper bound for z_0 is 2.42. This gives an edge length ratio of .992 and a hyper volume reduction of $(.992)^{**10} = .921$.

C. Summary

In part a. of this section it was shown that the bivariate/trivariate methods gave better upper bounds with respect to those given by the conservative assumption of independence as absolute values of correlations increased, number of X_i decreased and z became neither too large nor too close to zero. In part b it was shown that for large absolute correlations between X_{ij}, $e_{ij} \varepsilon T$ the bivariate method gave considerably lower upper bounds for $\alpha_0 | z_F$ and $z_0 | \alpha_F$ than did independence. It was also shown that if $P_{ij'}$ was large for $e_{ij,jj'} \varepsilon T'$ that the trivariate method gave considerably lower upper bounds for

$\alpha_0|z_F$ and $z_0|\alpha_F$ than did the bivariate method. If $P_{ij'}$ was small compared to $P_{ij'}$ i.e. $P_{ij'} = (P_{ij})^2$ then the upper bounds for $\alpha_0|z_F$ and $z_F|\alpha_0$ given by the trivariate method, were not much smaller than those given by the bivariate method. It is felt that the improvements in the upper bounds were enough to warrant usage of these methods in practice.

VI. Areas of Application.

Many areas of application of these techniques come to mind. Among them

a. Multiple Regression. Using the general linear model

$Y = XB + \varepsilon, \varepsilon$ iid $N(0,\sigma^2)$

σ^2 known, B an unknown $(n\times 1)$ vector

$\hat{B} = (X'X)^{-1}X'Y$

$Cov(\hat{B}) = (X'X)^{-1}\sigma^2$

corr $\hat{B} = W'(X'X)^{-1}W$ where W is an $(n\times 1)$ vector such that

$W_i = ([(X'X)^{-1}]_{ii})^{-1/2}$.

It is desired to obtain confidence intervals for B_i where $i = 1,\ldots,n$. The current method used to find the largest z_0 such that

$$(21) \qquad \Pr\left[\bigcup_{i=1}^{n} B_i \notin (\hat{B}_i - z_0(X'X)^{-1}_{ii}\sigma,\ \hat{B}_i + z_0(X'X)^{-1}_{ii}\sigma)\right] \le \alpha_F \,.$$

is the standard Bonferroni method which involves finding z_0 such that $\Pr[B_i \notin (\hat{B}_i - z_0(X'X)^{-1}_{ii}\sigma,\ \hat{B}_i + z_0(X'X)^{-1}_{ii}\sigma)] = \alpha_F/n$. Smaller values of z_0 which will have property (21) can be obtained by using (8) to form the "Best Bivariate Tree: T_0" then iteratively using the Newton-Raphson procedure (9.5) with $f(z) = (9)-\alpha_F$ and $f'(z_i) = (10)$. Even smaller values for z_0 could be obtained by applying (16) in some fashion to the tree T_0' created from T_0 by (13).

It is important to note that often σ is not known and is estimated by $\hat{\sigma} = \frac{(Y_i - \hat{Y})^2}{(n-1)}$ which then means that confidence intervals for B_i are created using $\hat{\sigma}$ instead of σ. All of the methods

described previously could then be easily modified to use multivariate t distributions instead of multivariate normal distributions.

b. Time Series.

These simultaneous inference procedures can be applied to estimations of probabilities of events involving autoregressive, moving average and other types of time series models. For instance let's look at the MA(3) case where

$Y_i = g_1(X_i)+g_2(X_{i-1})+g_3(X_i-2)$ where i = time; g_1,g_2,g_3 are known constants, and the X_i are iid $N(0,\sigma^2),\sigma^2$ known.

The goal is to produce simultaneous at least $(1-\alpha_F)$ confidence intervals for the values of $[Y_i,Y_{i+1},\ldots,Y_{i+n-1}]$ given no prior knowledge of $X_1,\ldots,X_i$ or $Y_1,\ldots,Y_{i-1}$. Without loss of generality let $\sigma^2 = 1$ and $[g_1^2+g_2^2+g_3^2] = 1$. Then

$$\Sigma = \begin{vmatrix} 1 & P_1 & P_2 & 0 & \cdots & 0 \\ P_1 & \ddots & \ddots & \ddots & \ddots & \vdots \\ P_2 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & P_2 \\ \vdots & \ddots & \ddots & \ddots & \ddots & P_1 \\ 0 & \cdots & 0 & P_2 & P_1 & 1 \end{vmatrix} \quad \text{where} \quad \begin{array}{l} P_1 = [g_1g_2+g_2g_3] \\ P_2 = [g_1g_3] \end{array} .$$

Using the notation $A = [\bigcup_{i=1}^{n} X_i \notin (-z,z)]$ and $[A_i = X_i \notin (-z,z)]$ then the best bivariate tree T_0 is $[A_i-A_{i+1}-\cdots-A_{i+q}]$. The standard Bonferroni z_0 is the largest z such that $\Pr(X_i \notin (-z_0,z_0) \leq \frac{\alpha}{n}$. Smaller values for z_0 could be computed iteratively by using the Newton-Raphson procedure (9.5) on T_0 with $f = (9)-\alpha_F$ and $f' = (10)$. Still smaller values for z_0 could be computer by applying (16) in some fashion to the tree T_0' created from T_0 by (13).

c. Linear Combinations of Variables.

Let $[X_1,X_2,\ldots,X_n]$ be iid $N(0,1)$. Let $a_1,a_2,\ldots,a_m$ be $(n\times 1)$ vectors and suppose one wishes to determine $(1-\alpha)$ simultaneous confidence intervals for $[a_1X,a_2X,\ldots,a_mX]$. If $n \leq 4$ and $a_1,\ldots,a_m$ are contrasts or if $n \leq 3$ then Uusipaika has developed a way to determine z_0 such that

$$\Pr\left\{\bigcup_{i=1}^{n} a_iX \notin (-z_0 \sum_{j=1}^{n} a_{ij}^2,\ z_0 \sum_{j=1}^{n} a_{ij}^2)\right\} = 1-\alpha_F .$$

For other cases, Tukey's method on the standard Bonferroni are used to find confidence intervals for these linear combinations of parameters which are at least $(1-\alpha_F)$ simultaneously confident. Letting $A_i = a_iX\notin(-z \sum_j a_{ij}^2, z\sum_j a_{ij}^2)$; (8) can be used to obtain the best bivariate tree: T_0 and the Newton-Raphson procedure (9.5) with $f(z)=(9)-\alpha_F$ and $f'(z)=(10)$ can be used to find a z_0 such that

$$\Pr\left\{\bigcup_{i=1}^{n}\left[a_iX \notin (-z_0 \sum_{j=1}^{m} a_i^2,\ z_0 \sum_{j=1}^{m} a_{ij}^2)\right]\right\} \geq (1-\alpha_F) .$$

Even smaller values for z_0 could be obtained by applying (16) in some fashion to the tree T_0' created from T_0 by (13). In this case we might need Uusipaika's formula to integrate over singular three-dimensional distributions since we are dealing with linear combinations which could mean singularity in the covariance matrices of three adjacent points along T_0.

d. Survival Times. (A special application).

Up until now the exact correlation structure of x_i, x_j, $i,j \varepsilon T_0$ has been assumed to be known. Below is presented an application of the bivariate method in a situation where the exact correlations are not known. The application will be quite complex and given without rigorous justification. The author plans to give more justification in a later technical report or thesis. Those who are not interested in this application may wish to skip this section.

Assume the classical survival model with no censoring where the time of death is a random variable. Let S_{t_i} be the probability of surviving to time t_i where $0 < t_1 < t_2 < \cdots < t_k$. Assume that a sample of m individuals is followed starting at time zero and the time of death of each individual is recorded. Let $\hat{S}_{t_1}$ = (number of individuals alive at time t_i)/m be the sample estimate of S_{t_i} for $i = 1,2,\ldots,k$. If m is large then by the multivariate central limit theorem

$$\sqrt{n} \begin{pmatrix} S_{t_1} - \hat{S}_{t_1} \\ S_{t_2} - \hat{S}_{t_2} \\ \vdots \\ S_{t_k} - \hat{S}_{t_k} \end{pmatrix} \underset{\text{approximately}}{\sim} N\left(\begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix} , \begin{pmatrix} \Sigma \end{pmatrix} \right)$$

where $\Sigma_{ii} = S_{t_i}(1-S_{t_i})$

and

$\Sigma_{ij} = (1-S_{t_i})S_{t_j}$ for $i < j$.

Often it is desirable to build $(1-\alpha_F)$ joint confidence intervals for the S_{t_i}, $i=1,2,\ldots,k$. Unfortunately, the variance of $(S_{t_i}-\hat{S}_{t_i})$ and the correlation of $(S_{t_i}-\hat{S}_{t_i}, S_{t_j}-\hat{S}_{t_j})$ depend on the true unknown values of S_{t_i} and S_{t_j}.

The current way to construct these simultaneous confidence intervals uses the Bonferroni method. For each time t_i, $i=1,\ldots,k$, upper and lower limits for the confidence interval of S_{t_i} are derived in (Fleiss page 14) and given below

(22a) $$S_{t_{u,i}} = \frac{B_i+\phi_i}{\xi_i} \quad : (S_{t_{u,i}} \text{ is the upper limit})$$

(22b) $$S_{t_{\ell,i}} = \frac{B_i-\phi_i}{\xi_i} \quad : (S_{t_{\ell,i}} \text{ is the lower limit})$$

where

$$B_i = 2m\hat{S}_{T_i} + (z_{(\alpha_{F/2k})})^2-1$$

$$\phi_i = z_{(\alpha_{F/2k})} \cdot \sqrt{(z_{(\alpha_{F/2k})})^2-(2+\tfrac{1}{m}) + 4\hat{S}_{t_i}(m(1-\hat{S}_{t_i})+1)}$$

$$\xi_i = 2(m+(z_{(\alpha_{F/2k})})^2)$$

and $z_{(\alpha_{F/2k})}$ is the value such that $\int_{-\infty}^{z_{(\alpha_{F/2k})}} \phi(x)dx = (1-\frac{\alpha_F}{2k})$

The $z_{(\alpha_{F/2k})}$ term in the above formulas was derived from the Bonferroni procedure: dividing $\alpha_{F/2}$ by k. Below is a procedure based on the bivariate method which derives a smaller value for the z term producing tighter intervals which will have at least a $(1-\alpha_F)$ joint confidence.

STEP I.

Clearly the best tree T_0 for this problem is $S_{t_1}-S_{t_2}-\cdots-S_{t_k}$. The sample estimate of the correlation between $\hat{S}_{t_i}$ and $\hat{S}_{t_{i+1}}$ is

$$\text{(23)} \qquad \mathrm{Cor}(\hat{S}_{t_i},\hat{S}_{t_{i+1}}) = \frac{\hat{S}_{t_{i+1}}(1-\hat{S}_{t_i})}{\sqrt{\hat{S}_{t_{i+1}}(1-\hat{S}_{t_{i+1}})\,\hat{S}_{t_i}(1-\hat{S}_{t_i})}}$$

$$\text{for } i=1,2,\ldots,k-1 .$$

So using these correlation estimates, derive $z_0|\alpha_F$ from the bivariate method applied to T_0 using the Newton-Raphson procedure with $f(z) = (9)-\alpha_F$ and $f'(z) = (10)$. Replace $z_{(\alpha_{F/2k})}$ by z_0 in equations (22a) and (22b) and let the k upper limits obtained be $S'_{t_{1u}},S'_{t_{2u}},\ldots,S'_{t_{ku}}$ and the k lower limits obtained by $S'_{t_{1\ell}},S'_{2_{2\ell}},\ldots,S'_{t_{k\ell}}$.

STEP II.

(a) Based on the values $S'_{t_{1u}},S'_{t_{2u}},\ldots,S'_{t_{ku}}$ calculate $\widehat{\mathrm{cor}}(S'_{t_{iu}},S'_{t_{i+1,u}})$ using formula (23). Then using these calculations, derive $z_0|\alpha_F$ from the bivariate method applied to T_0 using the Newton-Raphson procedure with $f(z) = (9)-\alpha_F$ and $f'(z) = (10)$. Replace $z_{(\alpha_{F/2k})}$ by this value of z_0 in equation (22a) to obtain new upper limits $S''_{t_{iu}},S''_{t_{2u}},\ldots,S''_{t_{ku}}$.

(b) Based on the values $S'_{t_{1\ell}}, S'_{t_{2\ell}}, \ldots, S'_{t_{k\ell}}$ calculate $\hat{cor}(S'_{t_{i\ell}}, S'_{t_{i+1,\ell}})$ using formula (23). Then using these correlations, derive $z_0 | \alpha_F$ from the bivariate method applied to T_0 using the Newton-Raphson procedure with $f(z) = (9) - \alpha_F$ and $f'(z) = (10)$. Replace $z_{(\alpha_{F/2k})}$ by this value of z_0 in equation (22b) to obtain new lower limits $S''_{t_{1\ell}}, S''_{t_{2\ell}}, \ldots, S''_{t_{k\ell}}$.

It is claimed that the $(S''_{t_{i\ell}}, S''_{t_{iu}})$ will form a set of joint confidence intervals of level at least $(1-\alpha_F)$ and that these intervals will be tighter than those produced using the Bonferroni method.

The justification for this claim is that in Step I, we calculate the upper limits $S'_{t_{1u}}, \ldots, S'_{t_{ku}}$ and use these upper limits to get a better estimate of the true upper bound correlation structure and with this, get the estimates $S''_{t_{1u}}, \ldots, S''_{t_{ku}}$ in Step II. Likewise for lower limits.

VII. Modifications on Regions, Distributions and Methods.

The improved Bonferroni method described which used integration over univariate, bivariate and trivariate distributions to get upper bounds for probabilities over multivariate normal cubes centered at zero could be slightly modified to handle the following situations.

A. Everything described previously could be applied to handle multivariate t distributions by replacing the univariate, bivariate and trivariate normal integrals by the appropriate univariate, bivariate and trivariate t integrals. The conservative assumption of independence, as an upper bound for the area outside the rectangle will, however, not necessarily hold for t-distribution regions. Also the correlation coefficients $P_{k,\ell}$ must be known (as will be the case for linear regression estimates of coefficients). Further study needs to be done for the case where $P_{k,\ell}$ are estimated.

B. The cubes centered at zero could be replaced with rectangular regions centered anywhere. Analogs of (9) and (16) could then be used replacing the lower limit of integration $-z$ with z_{Li} and the upper limit of integration z with z_{ui} for the i-th marginal. Unless the rectangle is centered exactly at the origin, the conservative assumption of independence as an upper bound for the area outside of the rectangle will not necessarily hold.

C. Some or all of the rectangular regions may be replaced with half plane or orthant like regions in which case analogs of (9) and (16) are used with either the lower limit of integration $-z$ being replaced by $-\infty$ or the upper limit of integration z being replaced by ∞. Again the conservative assumption of independence as an upper bound for the area outside of the region will not necessarily hold.

D. These methods could be modified to work for unions of disjoint sets of the types in B and C with normal or t distributions. The calculations might be quite complicated however.

VIII. References

Dunn, O.J. Tables of the Multivariate t Distributions, UCLA School of Public Health, 1968.

Dunnett, C.W., "Approximations to the probability integral and certain percentage points of the multivariate t distributions," Biometrika, 42, 258-260, 1955.

Fleiss, J. Statistical Methods for Rates and Proportions, John Wiley and Sons, New York, N.Y., 1981.

Hahn, G.J. and Henderson, R.M., "A table of percentage points of k student t variates and its applications," Biometrika, 58, 323-332, 1971.

Hoover, D. "Different algorithms for obtaining upper bounds to multivariate normal areas outside of origin centered rectangles using joint marginal distributions." Stanford University, Statistics Department Technical Report (1987).

Hunter, D., "An upper bound for the probability of a union," J. Appl. Prob., 13, 597-603, 1976.

Iyengar, S. "On evaluation of certain multivariate normal probabilities," Technical Report No. 322, Stanford University Statistics Department, 1982.

Jodego, K., "A simple proof of an inequality for multivariate normal probabilities of rectangles," Ann. Math. Statist., 41, 1357-1359, 1977.

Sidak, Z., "On multivariate normal probabilities of rectangles and their dependence on correlations," Ann. Math. Statist., 39, 1425-1434, 1968.

Uusipaika, E., "Exact simultaneous confidence intervals for multiple comparisons among three or four mean values," J. Amer. Statist. Assoc., 80, 196-201, 1985.

Worsley, K.J., "An improved Bonferroni inequality and applications," Biometrika, 69, 297-302, 1982.

DIRECTORY OF TABLES 1

Upper and Exact Bounds for $\alpha_0 = \Pr \bigcup_{i=1}^{n} X_i \notin (-\sigma_i z_F, \sigma_i z_F)$ where $X \sim N(0,\Sigma)$

for $z_F = 1.96,\ 2.25,\ 2.5,\ 2.75$

$n = 4,5,6,\ldots,20$

Calculated under:

Column (3): Conservative assumption of independence

Column (4): Using bivariate method assuming there exists a tree T connecting $(X_1, X_2, \ldots, X_n)$ with $|P_{x_i x_j}| = P$ for all $e_{ij}, \varepsilon T$.

Column (5): Using trivariate method assuming there exists a tree T as described for column 4 and a tree T' constructed from that tree such that $P_{x_i x_j} = P$ for all $e_{ij,jj'} \varepsilon T'$.

Column (6): The exact probability calculated under the assumption of equicorrelation $P_{x_k x_\ell} \equiv P$ for all $k;\ell$

Colmun (7): Using trivariate method assuming there exists a tree T as described in column 4 and a tree T' constructed from that tree such that $P_{x_i x_j} = P^2$ for all $e_{ij}, e_{jj'} \varepsilon T'$.

Table 1a. Upper Bounds For $\Pr\{(\bigcup_{i=1}^{n} X_i) \notin (-z,z)\}$

z=1.96

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj' \varepsilon T$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij'} = (P_{ij})^2$ $ij,jj' \varepsilon T'$
.100	3	0.14261	---	0.14188	0.14188	0.14212
.100	4	0.18548	---	---	0.18412	---
.100	5	0.22620	---	---	0.22408	---
.100	6	0.26489	---	---	0.26190	---
.100	7	0.30164	---	---	0.29771	---
.100	8	0.33656	---	---	0.33165	---
.100	9	0.36973	---	---	0.36382	---
.100	10	0.40124	---	---	0.39432	---
.100	11	0.43117	---	---	0.42325	---
.100	12	0.45961	---	---	0.45069	---
.100	13	0.48663	---	---	0.47674	---
.100	14	0.51299	---	---	0.50147	---
.100	15	0.53668	---	---	0.52496	---
.100	16	0.55984	---	---	0.54727	---
.100	17	0.58185	---	---	0.56846	---
.100	18	0.60275	---	---	0.58861	---
.100	19	0.62261	---	---	0.60775	---
.100	20	0.64148	---	---	0.62597	---
.200	3	0.14261	---	0.13972	0.13972	0.14059
.200	4	0.18548	---	0.18300	0.18047	0.18475
.200	5	0.22620	---	---	0.21827	---
.200	6	0.26489	---	---	0.25379	---
.200	7	0.30164	---	---	0.28721	---
.200	8	0.33656	---	---	0.31871	---
.200	9	0.36973	---	---	0.34845	---
.200	10	0.40124	---	---	0.37656	---
.200	11	0.43117	---	---	0.40315	---
.200	12	0.45961	---	---	0.42835	---
.200	13	0.48663	---	---	0.45224	---
.200	14	0.51229	---	---	0.47493	---
.200	15	0.53668	---	---	0.49648	---
.200	16	0.55984	---	---	0.51697	---
.200	17	0.58185	---	---	0.53647	---
.200	18	0.60275	---	---	0.55039	---
.200	19	0.62261	---	---	0.57274	---
.200	20	0.64148	---	---	0.58961	---

"---" means value is higher than that given by conservative assumption of independence.

Table 1a Continued.

z=1.96

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj'\epsilon T$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij'} = (P_{ij})^2$ $ij,jj'\epsilon T'$
.300	3	0.14261	0.14023	0.13623	0.13623	0.13802
.300	4	0.18548	0.18535	0.17734	0.17409	0.18092
.300	5	0.22620	---	0.21846	0.20891	0.22382
.300	6	0.26489	---	0.25959	0.24127	---
.300	7	0.30164	---	0.30069	0.27146	---
.300	8	0.33656	---	---	0.29971	---
.300	9	0.36973	---	---	0.32624	---
.300	10	0.40124	---	---	0.35122	---
.300	11	0.43117	---	---	0.37479	---
.300	12	0.45961	---	---	0.39708	---
.300	13	0.48663	---	---	0.41821	---
.300	14	0.51229	---	---	0.43827	---
.300	15	0.53668	---	---	0.45733	---
.300	16	0.55984	---	---	0.47549	---
.300	17	0.58185	---	---	0.49280	---
.300	18	0.60275	---	---	0.50932	---
.300	19	0.62261	---	---	0.52511	---
.300	20	0.64148	---	---	0.54021	---
.400	3	0.14261	0.13646	0.13146	0.13146	0.13427
.400	4	0.15848	0.17970	0.16968	0.16582	0.17531
.400	5	0.22620	0.22293	0.20791	0.19658	0.21635
.400	6	0.26489	---	0.24614	0.22557	0.25740
.400	7	0.30164	---	0.28436	0.25194	0.29844
.400	8	0.33656	---	0.32259	0.27645	---
.400	9	0.36973	---	0.36082	0.29932	---
.400	10	0.40124	---	0.39905	0.32075	---
.400		0.43117	---	---	0.34092	---
.400	12	0.45961	---	---	0.35995	---
.400	13	0.48663	---	---	0.37796	---
.400	14	0.51229	---	---	0.39505	---
.400	15	0.53668	---	---	0.41130	---
.400	16	0.55984	---	---	0.42678	---
.400	17	0.58185	---	---	0.44156	---
.400	18	0.60275	---	---	0.45568	---
.400	19	0.62261	---	---	0.46921	---
.400	20	0.64148	---	---	0.48218	---

"---" means value is higher than that given by conservative assumption of independence.

Table 1a Continued.

z=1.96

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj' \varepsilon T$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij'} = (P_{ij})^2$ $ij,jj' \varepsilon T'$
.500	3	0.14261	0.13348	0.12543	0.12543	0.12922
.500	4	0.18548	0.17223	0.16013	0.15582	0.16769
.500	5	0.22620	0.21297	0.19482	0.18285	0.20617
.500	6	0.26489	0.25371	0.22952	0.20728	0.24465
.500	7	0.30164	0.29445	0.26421	0.22959	0.28313
.500	8	0.33656	0.33520	0.29890	0.25013	0.32160
.500	9	0.36973	---	0.33360	0.26917	0.36008
.500	10	0.40124	---	0.36829	0.28693	0.39856
.500	11	0.43117	---	0.40299	0.30356	---
.500	12	0.45961	---	0.43768	0.31920	---
.500	13	0.48663	---	0.47238	0.33397	---
.500	14	0.51229	---	0.50705	0.34796	---
.500	15	0.53668	---	---	0.36125	---
.500	16	0.55984	---	---	0.37390	---
.500	17	0.58185	---	---	0.38597	---
.500	18	0.60275	---	---	0.39751	---
.500	19	0.62261	---	---	0.40857	---
.500	20	0.64148	---	---	0.41918	---
.600	3	0.14261	0.12509	0.11809	0.11809	0.12263
.600	4	0.18548	0.16263	0.14864	0.14404	0.15771
.600	5	0.22620	0.20018	0.17919	0.16677	0.19280
.600	6	0.26489	0.23773	0.20974	0.18681	0.22788
.600	7	0.30164	0.27527	0.24029	0.20498	0.26297
.600	8	0.33656	0.31282	0.27084	0.22154	0.29806
.600	9	0.36973	0.35037	0.30139	0.23678	0.33314
.600	10	0.40124	0.38791	0.33194	0.25090	0.36823
.600	11	0.43117	0.42546	0.36249	0.26406	0.40331
.600	12	0.45961	---	0.39303	0.27639	0.43840
.600	13	0.48663	---	0.42358	0.28798	0.47349
.600	14	0.51229	---	0.45413	0.29893	0.50957
.600	15	0.53668	---	0.48468	0.30931	---
.600	16	0.55984	---	0.51523	0.31917	---
.600	17	0.58185	---	0.54578	0.32857	---
.600	18	0.60275	---	0.57633	0.33754	---
.600	19	0.62261	---	0.60688	0.34613	---
.600	20	0.64148	---	0.63743	0.35437	---

"---" means value is higher than that given by conservative assumption of independence.

Table 1a Continued.

z=1.96

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj'\varepsilon T'$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij'} = (P_{ij})^2$ $ij,jj'\varepsilon T'$
.700	3	0.14261	0.11689	0.10920	0.10920	0.11414
.700	4	0.18458	0.15034	0.13496	0.13035	0.14483
.700	5	0.22620	0.18379	0.16072	0.14839	0.17553
.700	6	0.26489	0.21724	0.18648	0.16416	0.20623
.700	7	0.30164	0.25069	0.21224	0.17820	0.23692
.700	8	0.33656	0.28414	0.23799	0.19087	0.26762
.700	9	0.36973	0.31759	0.26375	0.20243	0.29841
.700	10	0.40124	0.35104	0.28951	0.21306	0.32901
.700	11	0.43117	0.38449	0.31527	0.22292	0.35970
.700	12	0.45961	0.41794	0.34103	0.23211	0.39040
.700	13	0.48663	0.45139	0.36679	0.24071	0.42109
.700	14	0.51229	0.48484	0.39254	0.24881	0.45179
.700	15	0.53668	0.51828	0.41830	0.25646	0.48248
.700	16	0.55984	0.55173	0.44406	0.26371	0.51318
.700	17	0.58185	---	0.46982	0.27059	0.54387
.700	18	0.60275	---	0.49558	0.27716	0.57457
.700	19	0.62261	---	0.52134	0.28343	0.60526
.700	20	0.64148	---	0.54709	0.28943	0.63596
.800	3	0.14261	0.10611	0.09822	0.09822	0.10307
.800	4	0.18548	0.13416	0.11838	0.11413	0.12809
.800	5	0.22620	0.16222	0.13855	0.12734	0.15310
.800	6	0.26489	0.19028	0.15871	0.13867	0.17812
.800	7	0.30164	0.21833	0.17888	0.14861	0.20314
.800	8	0.33656	0.24639	0.19904	0.15747	0.22815
.800	9	0.36973	0.27444	0.21921	0.16548	0.25317
.800	10	0.40124	0.30250	0.23927	0.17279	0.27819
.800	11	0.43117	0.33056	0.25954	0.17952	0.30320
.800	12	0.45961	0.35861	0.27970	0.18576	0.32822
.800	13	0.48663	0.38667	0.29987	0.19167	0.35324
.800	14	0.51229	0.41472	0.32003	0.19701	0.37825
.800	15	0.53668	0.44278	0.34020	0.20213	0.40327
.800	16	0.55984	0.47084	0.36036	0.20697	0.42829
.800	17	0.58185	0.49889	0.38053	0.21155	0.45331
.800	18	0.60275	0.52695	0.40069	0.21591	0.47832
.800	19	0.62261	0.55500	0.42086	0.22006	0.50334
.800	20	0.64148	0.58306	0.44102	0.22402	0.52836

"---" means value is higher than that given by conservative assumption of independence.

Table 1a Continued.

z=1.96

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj'\varepsilon T'$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij'} = (P_{ij})^2$ $ij,jj'\varepsilon T'$
.900	3	0.14261	0.09070	0.08364	0.08364	0.08764
.900	4	0.18548	0.11105	0.09693	0.09359	0.10493
.900	5	0.22620	0.13141	0.11023	0.10157	0.12223
.900	6	0.26489	0.15176	0.12352	0.10825	0.13952
.900	7	0.30164	0.17211	0.13681	0.11399	0.15681
.900	8	0.33656	0.19246	0.15011	0.11904	0.17410
.900	9	0.36973	0.21282	0.16340	0.12354	0.19140
.900	10	0.40124	0.23317	0.17669	0.12761	0.20869
.900	11	0.43117	0.25352	0.18999	0.13132	0.22598
.900	12	0.45961	0.27387	0.20328	0.13473	0.24328
.900	13	0.48663	0.29423	0.21657	0.13788	0.26057
.900	14	0.51229	0.31458	0.22987	0.14082	0.27786
.900	15	0.53668	0.33493	0.24316	0.14357	0.29515
.900	16	0.55984	0.35528	0.25645	0.14615	0.31245
.900	17	0.58185	0.37564	0.26975	0.14859	0.32974
.900	18	0.60275	0.39599	0.28304	0.15090	0.34703
.900	19	0.62261	0.41634	0.29633	0.15309	0.36433
.900	20	0.64148	0.43669	0.30963	0.15517	0.38162
.990	3	0.14261	0.06315	0.06018	0.06018	*
.990	4	0.18548	0.06973	0.06378	0.06264	*
.990	5	0.22620	0.07631	0.06739	0.06449	*
.990	6	0.26489	0.08289	0.07099	0.06597	*
.990	7	0.30164	0.08947	0.07460	0.06720	*
.990	8	0.33656	0.09605	0.07820	0.06825	*
.990	9	0.36973	0.10263	0.08181	0.06917	*
.990	10	0.40123	0.10920	0.08541	0.06998	*
.990	11	0.43117	0.11578	0.08902	0.07071	*
.990	12	0.45961	0.12236	0.09262	0.07136	*
.990	13	0.48663	0.12894	0.09623	0.07196	*
.990	14	0.51229	0.13552	0.09982	0.07252	*
.990	15	0.53668	0.14210	0.10344	0.07303	*
.990	16	0.55984	0.14868	0.10704	0.07350	*
.990	17	0.58185	0.15526	0.11065	0.07395	*
.990	18	0.60275	0.16183	0.11425	0.07436	*
.990	19	0.62261	0.16841	0.11786	0.07475	*
.990	20	0.64148	0.17499	0.12146	0.07513	*

"*" means value is missing due to inability of IMSL subroutine DMLIN to integrate.

Table 1b. Upper bounds for $Pr\{(\bigcup_{i=1}^{n} X_i) \notin (-z,z)\}$

z=2.25

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj' \varepsilon T'$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij'} = (P_{ij})^2$ $ij,jj' \varepsilon T'$
.100	3	0.07157	0.07195	0.07128	0.07128	0.07137
.100	4	0.09427	---	---	0.09455	---
.100	5	0.11641	---	---	0.11625	---
.100	6	0.13801	---	---	0.13726	---
.100	7	0.15909	---	---	0.15789	---
.100	8	0.17965	---	---	0.17786	---
.100	9	0.19971	---	---	0.19730	---
.100	10	0.21927	---	---	0.21622	---
.100	11	0.23836	---	---	0.23464	---
.100	12	0.25698	---	---	0.25256	---
.100	13	0.27515	---	---	0.27001	---
.100	14	0.29287	---	---	0.28701	---
.100	15	0.31016	---	---	0.30357	---
.100	16	0.32702	---	---	0.31969	---
.100	17	0.34348	---	---	0.33540	---
.100	18	0.35953	---	---	0.35070	---
.100	19	0.37519	---	---	0.36561	---
.100	20	0.39046	---	---	0.38010	---
.200	3	0.07157	0.07132	0.07039	0.07039	0.07075
.200	4	0.09427	---	0.09290	0.09318	0.09362
.200	5	0.11641	---	0.11540	0.11380	---
.200	6	0.13801	---	0.13791	0.13370	---
.200	7	0.15909	---	---	0.15242	---
.200	8	0.17965	---	---	0.17151	---
.200	9	0.19971	---	---	0.18949	---
.200	10	0.21927	---	---	0.20691	---
.200	11	0.23836	---	---	0.22377	---
.200	12	0.25698	---	---	0.24012	---
.200	13	0.27515	---	---	0.25598	---
.200	14	0.29287	---	---	0.27136	---
.200	15	0.31016	---	---	0.28630	---
.200	16	0.32702	---	---	0.30080	---
.200	17	0.34348	---	---	0.31490	---
.200	18	0.35953	---	---	0.32861	---
.200	19	0.37519	---	---	0.34194	---
.200	20	0.39046	---	---	0.35491	---

*"---" means value is higher than that given by conservative assumption of independence.

Table 1b Continued.

z=2.25

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj'\epsilon T'$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij'} = (P_{ij})^2$ $ij,jj'\epsilon T'$
.300	3	0.07157	0.07025	0.06892	0.06892	0.06968
.300	4	0.09427	0.09315	0.09049	0.08572	0.09200
.300	5	0.11641	0.11606	0.11205	0.10886	0.11433
.300	6	0.13801	---	0.13362	0.12723	0.13665
.300	7	0.15909	---	0.15519	0.14478	0.15898
.300	8	0.17965	---	0.17676	0.16158	---
.300	9	0.19971	---	0.19833	0.17770	---
.300	10	0.21927	---	---	0.19318	---
.300	11	0.23836	---	---	0.20808	---
.300	12	0.25698	---	---	0.22243	---
.300	13	0.25715	---	---	0.23628	---
.300	14	0.29287	---	---	0.24966	---
.300	15	0.31016	---	---	0.26259	---
.300	16	0.32702	---	---	0.27511	---
.300	17	0.34348	---	---	0.28724	---
.300	18	0.35953	---	---	0.29900	---
.300	19	0.37519	0.43668	0.41401	0.31041	---
.300	20	0.39046	0.45958	0.43557	0.32150	---
.400	3	0.07157	0.06868	0.06684	0.06684	0.06806
.400	4	0.09427	0.09079	0.08711	0.08572	0.08955
.400	5	0.11641	0.11291	0.10738	0.10317	0.11105
.400	6	0.13801	0.13502	0.12766	0.11956	0.13254
.400	7	0.15909	0.15714	0.14793	0.13503	0.15404
.400	8	0.17965	0.17925	0.16820	0.14968	0.17553
.400	9	0.19971	---	0.18848	0.16359	0.19702
.400	10	0.21927	---	0.20875	0.17686	0.21852
.400	11	0.23836	---	0.22902	0.18953	---
.400	12	0.25698	---	0.24930	0.20167	---
.400	13	0.27515	---	0.26957	0.21332	---
.400	14	0.29287	---	0.28984	0.22452	---
.400	15	0.31016	---	0.31011	0.23530	---
.400	16	0.32702	---	---	0.24571	---
.400	17	0.34348	---	---	0.25575	---
.400	18	0.35953	---	---	0.26547	---
.400	19	0.37519	---	---	0.27487	---
.400	20	0.39046	---	---	0.28399	---

*"---" means value is higher than that given by conservative assumption of independence.

Table 1b Continued.

z=2.25

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj'\varepsilon T'$	Exact for Equicorrelated Case $P_{k\ell} = P$ all k,ℓ	Trivariate $P_{ij'} = (P_{ij})^2$ $ij,jj'\varepsilon T'$
.500	3	0.07157	0.06651	0.06410	0.06410	0.06579
.500	4	0.09427	0.08754	0.08272	0.08090	0.08610
.500	5	0.11641	0.10857	0.10134	0.09625	0.10641
.500	6	0.13801	0.12960	0.11996	0.11042	0.12672
.500	7	0.15909	0.15063	0.13858	0.12361	0.14703
.500	8	0.17965	0.17166	0.15720	0.13595	0.16734
.500	9	0.19971	0.19269	0.17582	0.14757	0.18765
.500	10	0.21927	0.21372	0.19444	0.15855	0.20796
.500	11	0.23836	0.23475	0.21306	0.16896	0.22827
.500	12	0.25698	0.25578	0.23168	0.17887	0.24858
.500	13	0.27515	---	0.25030	0.18833	0.26889
.500	14	0.29287	---	0.26892	0.19739	0.28921
.500	15	0.31016	---	0.28754	0.20607	0.30952
.500	16	0.32702	---	0.30616	0.21441	---
.500	17	0.34348	---	0.32478	0.22433	---
.500	18	0.35953	---	0.34340	0.23017	---
.500	19	0.37519	---	0.36202	0.23765	---
.500	20	0.39046	---	0.38064	0.24487	---
.600	3	0.07157	0.06359	0.06061	0.06061	0.06271
.600	4	0.09427	0.08316	0.07720	0.07513	0.08140
.600	5	0.11641	0.10273	0.09379	0.08811	0.10009
.600	6	0.13801	0.12230	0.11038	0.09934	0.11878
.600	7	0.15909	0.14188	0.12697	0.11068	0.13747
.600	8	0.17965	0.16145	0.14356	0.12067	0.15616
.600	9	0.19971	0.18102	0.16015	0.12998	0.17485
.600	10	0.21927	0.20059	0.17674	0.13870	0.19355
.600	11	0.23836	0.22016	0.19333	0.14692	0.21224
.600	12	0.25698	0.23973	0.20992	0.15469	0.23093
.600	13	0.27515	0.25930	0.22651	0.16207	0.24962
.600	14	0.29287	0.27887	0.24309	0.16909	0.26831
.600	15	0.31016	0.29844	0.25968	0.17580	0.28700
.600	16	0.32702	0.31802	0.27627	0.18221	0.30569
.600	17	0.34348	0.33759	0.29286	0.18837	0.32438
.600	18	0.35953	0.35716	0.30945	0.19428	0.34307
.600	19	0.37519	---	0.32604	0.19997	0.36176
.600	20	0.39046	---	0.34263	0.20547	0.38045

*"---" means value is higher than that given by conservative assumption of independence.

Table 1b Continued.

z=2.25

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj'\varepsilon T'$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij'} = (P_{ij})^2$ $ij,jj'\varepsilon T'$
.700	3	0.07157	0.05968	0.05620	0.05620	0.05858
.700	4	0.09427	0.07729	0.07033	0.06814	0.07511
.700	5	0.11641	0.09491	0.08447	0.07854	0.09163
.700	6	0.13801	0.11252	0.09860	0.08780	0.10815
.700	7	0.15909	0.13013	0.11274	0.09617	0.12467
.700	8	0.17965	0.14775	0.12687	0.10381	0.14119
.700	9	0.19971	0.16536	0.14101	0.11086	0.15771
.700	10	0.21927	0.18298	0.15514	0.11741	0.17424
.700	11	0.23836	0.20059	0.16927	0.12353	0.19076
.700	12	0.25698	0.21820	0.18341	0.12928	0.20728
.700	13	0.27515	0.23582	0.19754	0.13471	0.22380
.700	14	0.29287	0.25343	0.21168	0.13985	0.24032
.700	15	0.31016	0.27105	0.22581	0.14473	0.25684
.700	16	0.32702	0.28866	0.23995	0.14934	0.27337
.700	17	0.34348	0.30628	0.25408	0.15383	0.28989
.700	18	0.35953	0.32389	0.26822	0.15809	0.30641
.700	19	0.37519	0.34150	0.28235	0.16218	0.32293
.700	20	0.39046	0.35912	0.29649	0.16611	0.33945
.800	3	0.07157	0.05430	0.05053	0.05053	0.05299
.800	4	0.09427	0.06923	0.06169	0.05956	0.06659
.800	5	0.11641	0.08416	0.07284	0.06714	0.08020
.800	6	0.13801	0.09909	0.08399	0.07384	0.09381
.800	7	0.15909	0.11401	0.09515	0.07974	0.10742
.800	8	0.17965	0.12895	0.10630	0.08505	0.12103
.800	9	0.19971	0.14387	0.11746	0.08990	0.13464
.800	10	0.21927	0.15880	0.12861	0.09439	0.14825
.800	11	0.23836	0.17372	0.13977	0.09848	0.16186
.800	12	0.25698	0.18865	0.15092	0.10233	0.17547
.800	13	0.27515	0.20358	0.16208	0.10594	0.18908
.800	14	0.29287	0.21851	0.17323	0.10934	0.20268
.800	15	0.31016	0.23343	0.18439	0.11255	0.21629
.800	16	0.32702	0.24836	0.19554	0.11560	0.22990
.800	17	0.34348	0.26329	0.20670	0.11850	0.24351
.800	18	0.35953	0.27821	0.21785	0.12126	0.25712
.800	19	0.37519	0.29314	0.22900	0.12391	0.27073
.800	20	0.39046	0.30807	0.24016	0.12644	0.28434

Table 1b Continued.

z=2.25

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj'\varepsilon T'$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij'} = (P_{ij})^2$ $ij,jj'\varepsilon T'$
.900	3	0.07157	0.04633	0.04277	0.04277	0.04485
.900	4	0.09427	0.05727	0.05015	0.04840	0.05437
.900	5	0.11641	0.06821	0.05753	0.05299	0.06375
.900	6	0.13801	0.07915	0.06491	0.05686	0.07232
.900	7	0.15909	0.09010	0.07229	0.06022	0.08278
.900	8	0.17965	0.10104	0.07967	0.06319	0.09231
.900	9	0.19971	0.11198	0.08705	0.06586	0.10752
.900	10	0.21927	0.12292	0.09443	0.06820	0.11019
.900	11	0.23836	0.13386	0.10181	0.07051	0.12132
.900	12	0.25698	0.14480	0.10910	0.07256	0.13030
.900	13	0.27515	0.15574	0.11657	0.07447	0.13979
.900	14	0.29287	0.16668	0.12395	0.07625	0.14882
.900	15	0.31016	0.17762	0.13133	0.07792	0.15903
.900	16	0.32702	0.18856	0.13871	0.07950	0.16805
.900	17	0.34348	0.19951	0.14609	0.08099	0.17794
.900	18	0.35953	0.21045	0.15347	0.08241	0.18621
.900	19	0.37519	0.22139	0.16085	0.08375	0.19728
.900	20	0.39046	0.23233	0.16823	0.08504	0.20571
.990	3	0.07157	0.03159	0.03000	0.03000	*
.990	4	0.09427	0.03516	0.03198	0.03136	*
.990	5	0.11641	0.03873	0.03396	0.03234	*
.990	6	0.13801	0.04230	0.03594	0.03322	*
.990	7	0.15909	0.04586	0.03792	0.03391	*
.990	8	0.17965	0.04943	0.03990	0.03450	*
.990	9	0.19971	0.05300	0.04188	0.03511	*
.990	10	0.21927	0.05657	0.04387	0.03547	*
.990	11	0.23836	0.06014	0.04585	0.03588	*
.990	12	0.25698	0.06371	0.04783	0.03625	*
.990	13	0.27515	0.06728	0.04981	0.03659	*
.990	14	0.29287	0.07085	0.05179	0.03691	*
.990	15	0.31016	0.07442	0.05377	0.03719	*
.990	16	0.32702	0.07799	0.05575	0.03746	*
.990	17	0.34348	0.08156	0.05773	0.03771	*
.990	18	0.35935	0.08513	0.05971	0.03795	*
.990	19	0.37519	0.08870	0.06169	0.03817	*
.990	20	0.39046	0.09227	0.06368	0.03834	*

"*" means value is missing due to the inability of DMLIN to integrate.

Table 1c. Upper bounds for $\Pr\{(\bigcup_{i=1}^{n} X_i) \not\in (-z,z)\}$

z=2.50

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj' \varepsilon T'$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij'} = (P_{ij})^2$ $ij,jj' \varepsilon T'$
.100	3	0.03680	---	0.03669	0.03669	0.03672
.100	4	0.04876	---	0.04873	0.04854	---
.100	5	0.06057	---	---	0.06220	---
.100	6	0.07224	---	---	0.07170	---
.100	7	0.08376	---	---	0.08303	---
.100	8	0.09514	---	---	0.09418	---
.100	9	0.10638	---	---	0.10517	---
.100	10	0.11748	---	---	0.11599	---
.100	11	0.12844	---	---	0.12665	---
.100	12	0.13926	---	---	0.13716	---
.100	13	0.14995	---	---	0.14752	---
.100	14	0.16051	---	---	0.15772	---
.100	15	0.17093	---	---	0.16778	---
.100	16	0.18123	---	---	0.17769	---
.100	17	0.19140	---	---	0.18746	---
.100	18	0.20144	---	---	0.19710	---
.100	19	0.21136	---	---	0.20659	---
.100	20	0.22115	---	---	0.21595	---
.200	3	0.03680	0.03663	0.03634	0.03634	0.03648
.200	4	0.04876	0.04874	0.04815	0.04786	0.04844
.200	5	0.06057	---	0.05996	0.05913	0.06039
.200	6	0.07224	---	0.07177	0.07013	---
.200	7	0.08376	---	0.08359	0.09090	---
.200	8	0.09514	---	---	0.09143	---
.200	9	0.10638	---	---	0.10516	---
.200	10	0.11748	---	---	0.11599	---
.200	11	0.12844	---	---	0.12174	---
.200	12	0.13926	---	---	0.13144	---
.200	13	0.14995	---	---	0.14095	---
.200	14	0.16051	---	---	0.15028	---
.200	15	0.17093	---	---	0.15943	---
.200	16	0.18123	---	---	0.16288	---
.200	17	0.19140	---	---	0.17724	---
.200	18	0.20144	---	---	0.18590	---
.200	19	0.21136	---	---	0.19441	---
.200	20	0.22115	---	---	0.22781	---

*"---"means value is higher than that given by conservative assumption of independence.

Table 1c Continued.

z=2.50

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj' \varepsilon T'$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij'} = (P_{ij})^2$ $ij,jj' \varepsilon T'$
.300	3	0.03680	0.03620	0.03573	0.03573	0.03604
.300	4	0.04876	0.04809	0.04715	0.04673	0.04777
.300	5	0.06057	0.05999	0.05857	0.05733	0.05950
.300	6	0.07224	0.07188	0.06999	0.06758	0.07132
.300	7	0.08376	---	0.08142	0.07750	0.08296
.300	8	0.09514	---	0.09284	0.08712	0.09470
.300	9	0.10638	---	0.10426	0.09646	---
.300	10	0.11748	---	0.11568	0.10553	---
.300	11	0.12844	---	0.12710	0.11436	---
.300	12	0.13926	---	0.13852	0.12295	---
.300	13	0.14995	---	0.14994	0.13133	---
.300	14	0.16051	---	---	0.13950	---
.300	15	0.17093	---	---	0.14747	---
.300	16	0.18123	---	---	0.15526	---
.300	17	0.19140	---	---	0.16288	---
.300	18	0.20144	---	---	0.17032	---
.300	19	0.21136	---	---	0.17761	---
.300	20	0.22115	---	---	0.18475	---
.400	3	0.03680	0.03554	0.03483	0.03485	0.03535
.400	4	0.04876	0.04710	0.04568	0.04509	0.04672
.400	5	0.06057	0.05866	0.05653	0.05481	0.05809
.400	6	0.07224	0.07022	0.06738	0.06408	0.06946
.400	7	0.08376	0.08178	0.07824	0.07295	0.08083
.400	8	0.09514	0.09334	0.08909	0.08145	0.09220
.400	9	0.10638	0.10491	0.09994	0.08962	0.10357
.400	10	0.11748	0.11647	0.11079	0.09749	0.11494
.400	11	0.12844	0.12803	0.12164	0.10509	0.12631
.400	12	0.13926	---	0.13249	0.11244	0.13768
.400	13	0.14995	---	0.14334	0.11956	0.14905
.400	14	0.16051	---	0.15419	0.12645	0.16043
.400	15	0.17093	---	0.16504	0.13315	---
.400	16	0.18123	---	0.17590	0.13966	---
.400	17	0.19140	---	0.18675	0.14599	---
.400	18	0.20144	---	0.19760	0.15216	---
.400	19	0.21136	---	0.20845	0.15812	---
.400	20	0.22115	---	0.21930	0.16403	---

*"---" means value is higher than that given by conservative assumption of independence.

Table 1c Continued.

z=2.50

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj'\varepsilon T'$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij'} = (P_{ij})^2$ $ij,jj'\varepsilon T'$
.500	3	0.03680	0.03458	0.03358	0.03358	0.03433
.500	4	0.04876	0.04566	0.04367	0.04288	0.04517
.500	5	0.06057	0.05674	0.05375	0.05154	0.05600
.500	6	0.07224	0.06782	0.06383	0.05965	0.06684
.500	7	0.08376	0.07890	0.07391	0.06731	0.07767
.500	8	0.09514	0.08998	0.08400	0.07456	0.08851
.500	9	0.10638	0.10106	0.09408	0.08146	0.09934
.500	10	0.11748	0.11214	0.10416	0.08805	0.11018
.500	11	0.12844	0.12322	0.11425	0.09436	0.12101
.500	12	0.13926	0.13431	0.12433	0.10042	0.13184
.500	13	0.14995	0.14539	0.13441	0.10624	0.14268
.500	14	0.16051	0.15647	0.14450	0.11856	0.15351
.500	15	0.17093	0.16755	0.15458	0.12028	0.16435
.500	16	0.18123	0.17863	0.16466	0.12252	0.17518
.500	17	0.19140	0.18971	0.17475	0.12760	0.18602
.500	18	0.20144	0.20079	0.18483	0.13252	0.19685
.500	19	0.21136	---	0.19491	0.13706	0.20769
.500	20	0.22115	---	0.20499	0.14195	0.21852
.600	3	0.03680	0.03322	0.03191	0.03191	0.03289
.600	4	0.04876	0.04362	0.04100	0.04004	0.04296
.600	5	0.06057	0.05402	0.05008	0.04745	0.05302
.600	6	0.07224	0.06443	0.05917	0.05427	0.06309
.600	7	0.08376	0.07483	0.06826	0.06060	0.07316
.600	8	0.09514	0.08523	0.07734	0.06653	0.08323
.600	9	0.10638	0.09563	0.08643	0.07211	0.09330
.600	10	0.11748	0.10603	0.09552	0.07739	0.10337
.600	11	0.12844	0.11643	0.10461	0.08240	0.11343
.600	12	0.13926	0.12683	0.11369	0.08718	0.12350
.600	13	0.14995	0.13723	0.12278	0.09174	0.13357
.600	14	0.16051	0.14764	0.13187	0.09611	0.14364
.600	15	0.17093	0.15804	0.14095	0.10031	0.15371
.600	16	0.18123	0.16844	0.15004	0.10435	0.16377
.600	17	0.19140	0.17884	0.15913	0.10825	0.17384
.600	18	0.20144	0.18924	0.16822	0.11201	0.18391
.600	19	0.21136	0.19964	0.17730	0.11565	0.19398
.600	20	0.22115	0.21004	0.18639	0.11918	0.20405

*"---" means value is higher than that given by conservative assumption of independence.

Table 1c Continued.

z=2.50

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj' \varepsilon T'$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij'} = (P_{ij})^2$ $ij,jj' \varepsilon T'$
.700	3	0.03680	0.03132	0.02969	0.02969	0.03086
.700	4	0.04876	0.04076	0.03752	0.03645	0.03986
.700	5	0.06057	0.05021	0.04534	0.04244	0.04885
.700	6	0.07224	0.05966	0.05317	0.04785	0.05785
.700	7	0.08376	0.06911	0.06099	0.05279	0.06684
.700	8	0.09514	0.07855	0.06882	0.05735	0.07584
.700	9	0.10638	0.08800	0.07664	0.06159	0.08483
.700	10	0.11748	0.09745	0.08447	0.06557	0.09383
.700	11	0.12844	0.10690	0.09229	0.06930	0.10283
.700	12	0.13926	0.11635	0.10011	0.07284	0.11182
.700	13	0.14995	0.12579	0.10794	0.07620	0.12082
.700	14	0.16051	0.13524	0.11576	0.07939	0.12981
.700	15	0.17093	0.14469	0.12359	0.08244	0.13881
.700	16	0.18123	0.15415	0.13141	0.08536	0.14780
.700	17	0.19140	0.16359	0.13924	0.08817	0.15680
.700	18	0.20144	0.17303	0.14706	0.09087	0.16579
.700	19	0.21136	0.18248	0.15489	0.09346	0.17479
.700	20	0.22115	0.19193	0.16271	0.09597	0.18378
.800	3	0.03680	0.02859	0.02673	0.02673	0.02799
.800	4	0.04876	0.03667	0.03296	0.03188	0.03548
.800	5	0.06057	0.04475	0.03919	0.03629	0.04297
.800	6	0.07224	0.05284	0.04542	0.04018	0.05047
.800	7	0.08376	0.06092	0.05165	0.04367	0.05796
.800	8	0.09514	0.06900	0.05788	0.04684	0.06545
.800	9	0.10638	0.07709	0.06411	0.04988	0.07294
.800	10	0.11748	0.08517	0.07034	0.05244	0.08043
.800	11	0.12844	0.09325	0.07657	0.05500	0.08792
.800	12	0.13926	0.10134	0.08280	0.05731	0.09541
.800	13	0.14995	0.10942	0.08903	0.05952	0.10290
.800	14	0.16051	0.11750	0.09526	0.06162	0.11039
.800	15	0.17093	0.12559	0.10149	0.06361	0.11788
.800	16	0.18123	0.13367	0.10772	0.06550	0.12537
.800	17	0.19140	0.14175	0.11395	0.06731	0.13286
.800	18	0.20144	0.14984	0.12018	0.06904	0.14035
.800	19	0.21136	0.15792	0.12641	0.07070	0.14784
.800	20	0.22115	0.16600	0.13264	0.07224	0.15534

Table 1c Continued.

z=2.50

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj' \varepsilon T'$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij'} = (P_{ij})^2$ $ij,jj' \varepsilon T'$
.900	3	0.03680	0.02439	0.02255	0.02255	0.02368
.900	4	0.04876	0.03037	0.02669	0.02576	0.02895
.900	5	0.06057	0.03635	0.03083	0.02840	0.03422
.900	6	0.07224	0.04233	0.03498	0.03066	0.03950
.900	7	0.08376	0.04832	0.03912	0.03263	0.04477
.900	8	0.09514	0.05430	0.04326	0.03439	0.05004
.900	9	0.10638	0.06028	0.04741	0.03597	0.05531
.900	10	0.11748	0.06626	0.05155	0.03742	0.06059
.900	11	0.12844	0.07225	0.05569	0.03875	0.06586
.900	12	0.13926	0.07823	0.05984	0.03998	0.07113
.900	13	0.14995	0.08421	0.06398	0.04113	0.07641
.900	14	0.16051	0.09020	0.06813	0.04211	0.08168
.900	15	0.17093	0.09618	0.07227	0.04322	0.08695
.900	16	0.18123	0.10216	0.07641	0.04417	0.09223
.900	17	0.19140	0.10814	0.08056	0.04508	0.09750
.900	18	0.20144	0.11413	0.08470	0.04594	0.10277
.900	19	0.21136	0.12011	0.08884	0.04677	0.10805
.900	20	0.22115	0.12609	0.09299	0.04755	0.11332
.990	3	0.03680	0.01636	0.01549	0.01549	*
.990	4	0.04876	0.01833	0.01660	0.01626	*
.990	5	0.06057	0.02030	0.01771	0.01684	*
.990	6	0.07224	0.02227	0.01881	0.01730	*
.990	7	0.08376	0.02423	0.01991	0.01770	*
.990	8	0.09514	0.02620	0.02102	0.01803	*
.990	9	0.10638	0.02817	0.02212	0.01832	*
.990	10	0.11748	0.03014	0.02323	0.01858	*
.990	11	0.12844	0.03211	0.02433	0.01881	*
.990	12	0.13926	0.03408	0.02544	0.01903	*
.990	13	0.14995	0.03605	0.02654	0.01922	*
.990	14	0.16051	0.03802	0.02765	0.01940	*
.990	15	0.17093	0.03999	0.02875	0.01957	*
.990	16	0.18123	0.04196	0.02986	0.01972	*
.990	17	0.19140	0.04393	0.03096	0.01987	*
.990	18	0.20144	0.04590	0.03207	0.02000	*
.990	19	0.21136	0.04787	0.03317	0.02013	*
.990	20	0.22115	0.04983	0.03428	0.02025	*

* "*" means value is missing due to inability of DMLIN to integrate.

Table 1d. Upper Bounds for $\Pr\{(\bigcup_{i=1}^{n} X_i) \not\subset (-z,z)\}$

z=2.75

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj'\varepsilon T'$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij} = (P_{ij})^2$ $ij,jj'\varepsilon T'$
.100	3	0.01777	---	0.01774	0.01774	0.01775
.100	4	0.02363	---	0.02361	0.02359	0.02363
.100	5	0.02944	---	---	0.02945	---
.100	6	0.03523	---	---	0.03517	---
.100	7	0.04098	---	---	0.04084	---
.100	8	0.04669	---	---	0.04647	---
.100	9	0.05237	---	---	0.05205	---
.100	10	0.05802	---	---	0.05759	---
.100	11	0.06364	---	---	0.06309	---
.100	12	0.06922	---	---	0.06855	---
.100	13	0.07476	---	---	0.07396	---
.100	14	0.08028	---	---	0.07933	---
.100	15	0.08576	---	---	0.08467	---
.100	16	0.09121	---	---	0.08995	---
.100	17	0.09662	---	---	0.09521	---
.100	18	0.10201	---	---	0.10042	---
.100	19	0.10736	---	---	0.10559	---
.100	20	0.11268	---	---	0.11073	---
.200	3	0.01777	0.01770	0.01762	0.01762	0.01767
.200	4	0.02363	0.02357	0.02340	0.02341	0.02350
.200	5	0.02944	0.02944	0.02919	0.02945	0.02934
.200	6	0.03523	---	0.03498	0.03455	0.03518
.200	7	0.04098	---	0.04077	0.04002	---
.200	8	0.04669	---	0.04655	0.04541	---
.200	9	0.05237	---	0.05234	0.05073	---
.200	10	0.05802	---	---	0.05599	---
.200	11	0.06364	---	---	0.06118	---
.200	12	0.06922	---	---	0.06630	---
.200	13	0.07476	---	---	0.07137	---
.200	14	0.08028	---	---	0.07638	---
.200	15	0.08576	---	---	0.08132	---
.200	16	0.09121	---	---	0.08621	---
.200	17	0.09662	---	---	0.09105	---
.200	18	0.10201	---	---	0.09583	---
.200	19	0.10736	---	---	0.10056	---
.200	20	0.11268	---	---	0.10524	---

*"---" means value is higher than that given by conservative assumption of independence.

Table 1d Continued.

z=2.75

RHO	Number	Independence	Bivariate	Trivariate $P_{ij} = P_{ij}$ $ij,jj'\varepsilon T'$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij} = (P_{ij})^2$ $ij,jj'\varepsilon T'$
.300	3	0.01777	0.01755	0.01740	0.01740	0.01751
.300	4	0.02363	0.02334	0.02304	0.02295	0.02326
.300	5	0.02944	0.02913	0.02868	0.02831	0.02901
.300	6	0.03523	0.03493	0.03432	0.03354	0.03477
.300	7	0.04098	0.04072	0.03996	0.03867	0.04052
.300	8	0.04669	0.04652	0.04561	0.04368	0.04628
.300	9	0.05237	0.05231	0.05125	0.04860	0.05203
.300	10	0.05802	---	0.05689	0.05342	0.05778
.300	11	0.06364	---	0.06253	0.05815	0.06354
.300	12	0.06922	---	0.06818	0.06279	---
.300	13	0.07476	---	0.07382	0.06735	---
.300	14	0.08028	---	0.07946	0.07183	---
.300	15	0.08576	---	0.08510	0.07623	---
.300	16	0.09121	---	0.09075	0.08056	---
.300	17	0.09662	---	0.09639	0.08482	---
.300	18	0.10201	---	---	0.08902	---
.300	19	0.10736	---	---	0.09315	---
.300	20	0.11268	---	---	0.09721	---
.400	3	0.01777	0.01730	0.01704	0.01704	0.01724
.400	4	0.02363	0.02296	0.02246	0.02227	0.02286
.400	5	0.02944	0.02863	0.02788	0.02726	0.02847
.400	6	0.03523	0.03430	0.03329	0.03208	0.03408
.400	7	0.04098	0.03997	0.03871	0.03675	0.03970
.400	8	0.04669	0.04563	0.04413	0.04127	0.04531
.400	9	0.05237	0.05130	0.04955	0.04567	0.05093
.400	10	0.05802	0.05697	0.05496	0.04993	0.05654
.400	11	0.06364	0.06264	0.06038	0.05409	0.06215
.400	12	0.06922	0.06831	0.06580	0.05813	0.06777
.400	13	0.07476	0.07397	0.07121	0.06208	0.07338
.400	14	0.08028	0.07964	0.07663	0.06594	0.07900
.400	15	0.08576	0.08531	0.08205	0.06971	0.08461
.400	16	0.09121	0.09098	0.08746	0.07340	0.09022
.400	17	0.09662	---	0.09288	0.07701	0.09584
.400	18	0.10201	---	0.09830	0.08054	0.10145
.400	19	0.10736	---	0.10371	0.08401	0.10707
.400	20	0.11268	---	0.10913	0.08740	0.11268

*"--" Means value is higher than that given by conservative assumption of independence.

Table 1d Continued.

z=2.75

RHO	Number	Independence	Bivariate	Trivariate $P_{ij'} = P_{ij}$ $ij,jj'\epsilon T'$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij'} = (P_{ij})^2$ $ij,jj'\epsilon T'$
.500	3	0.01777	0.01691	0.01653	0.01653	0.01683
.500	4	0.02363	0.02238	0.02162	0.02131	0.02222
.500	5	0.02944	0.02785	0.02671	0.02583	0.02762
.500	6	0.03523	0.03333	0.03180	0.03013	0.03302
.500	7	0.04098	0.03880	0.03690	0.03423	0.03841
.500	8	0.04669	0.04427	0.04199	0.03815	0.04381
.500	9	0.05237	0.04974	0.04708	0.04193	0.04920
.500	10	0.05802	0.05522	0.05218	0.04556	0.05460
.500	11	0.06364	0.06069	0.05727	0.04907	0.05999
.500	12	0.06922	0.06616	0.06236	0.05246	0.06539
.500	13	0.07476	0.07164	0.06745	0.05575	0.07079
.500	14	0.08028	0.07711	0.07255	0.05893	0.07618
.500	15	0.08576	0.08258	0.07764	0.06203	0.08158
.500	16	0.09121	0.08806	0.08273	0.06504	0.08697
.500	17	0.09662	0.09353	0.08783	0.06798	0.09237
.500	18	0.10201	0.09900	0.09292	0.07084	0.09777
.500	19	0.10736	0.10448	0.09801	0.07363	0.10316
.500	20	0.11268	0.10995	0.10310	0.07635	0.10856
.600	3	0.01777	0.01632	0.01579	0.01579	0.01621
.600	4	0.02363	0.02150	0.02043	0.02003	0.02127
.600	5	0.02944	0.02669	0.02508	0.02395	0.02634
.600	6	0.03523	0.03187	0.02972	0.02761	0.03140
.600	7	0.04098	0.03705	0.03437	0.03106	0.03647
.600	8	0.04669	0.04223	0.03901	0.03431	0.04153
.600	9	0.05237	0.04741	0.04366	0.03740	0.04660
.600	10	0.05802	0.05259	0.04830	0.04035	0.05166
.600	11	0.06364	0.05778	0.05295	0.04317	0.05673
.600	12	0.06922	0.06296	0.05759	0.04587	0.06179
.600	13	0.07476	0.06814	0.06224	0.04848	0.06686
.600	14	0.08028	0.07332	0.06688	0.05098	0.07193
.600	15	0.08576	0.07850	0.07153	0.05341	0.07699
.600	16	0.09121	0.08368	0.07617	0.05575	0.08206
.600	17	0.09662	0.08887	0.08082	0.05802	0.08712
.600	18	0.10201	0.09405	0.08546	0.06022	0.09219
.600	19	0.10736	0.09923	0.09011	0.06236	0.09725
.600	20	0.11268	0.10441	0.09475	0.06444	0.10232

Table 1d Continued.

z=2.75

RHO	Number	Independence	Bivariate	Trivariate $P_{ij} = P_{ij}$ $ij,jj'\varepsilon T'$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij} = (P_{ij})^2$ $ij,jj'\varepsilon T'$
.700	3	0.01777	0.01546	0.01476	0.01476	0.01529
.700	4	0.02363	0.02021	0.01880	0.01832	0.01987
.700	5	0.02944	0.02496	0.02285	0.02154	0.02444
.700	6	0.03523	0.02971	0.02690	0.02447	0.02902
.700	7	0.04098	0.03446	0.03094	0.02718	0.03360
.700	8	0.04669	0.03921	0.03499	0.02971	0.03818
.700	9	0.05237	0.04396	0.03904	0.03208	0.04275
.700	10	0.05802	0.04872	0.04308	0.03431	0.04733
.700	11	0.06364	0.05347	0.04713	0.03643	0.05191
.700	12	0.06922	0.05822	0.05118	0.03845	0.05649
.700	13	0.07476	0.06297	0.05522	0.04036	0.06106
.700	14	0.08028	0.06772	0.05927	0.04221	0.06564
.700	15	0.08576	0.07247	0.06332	0.04397	0.07022
.700	16	0.09121	0.07722	0.06736	0.04567	0.07480
.700	17	0.09662	0.08197	0.07141	0.04730	0.07937
.700	18	0.10201	0.08672	0.07546	0.04888	0.08395
.700	19	0.10736	0.09145	0.07950	0.05041	0.08853
.700	20	0.11268	0.09622	0.08355	0.05189	0.09311
.800	3	0.01777	0.01417	0.01332	0.01332	0.01392
.800	4	0.02363	0.01828	0.01657	0.01606	0.01778
.800	5	0.02944	0.02238	0.01983	0.01844	0.02164
.800	6	0.03523	0.02649	0.02309	0.02057	0.02550
.800	7	0.04098	0.03059	0.02634	0.02249	0.02935
.800	8	0.04669	0.03470	0.02960	0.02426	0.03321
.800	9	0.05237	0.03880	0.03285	0.02588	0.03707
.800	10	0.05802	0.04291	0.03611	0.02740	0.04093
.800	11	0.06364	0.04701	0.03936	0.02882	0.04478
.800	12	0.06922	0.05112	0.04262	0.03076	0.04864
.800	13	0.07476	0.05522	0.04587	0.03143	0.05250
.800	14	0.08028	0.05933	0.04913	0.03263	0.05636
.800	15	0.08576	0.06343	0.05238	0.03378	0.06022
.800	16	0.09121	0.06754	0.05564	0.03487	0.06407
.800	17	0.09662	0.07164	0.05889	0.03592	0.06793
.800	18	0.10201	0.07575	0.06215	0.03692	0.07179
.800	19	0.10736	0.07986	0.06540	0.03789	0.07565
.800	20	0.11268	0.08396	0.06866	0.03882	0.07950

Table 1d Continued.

z=2.75

RHO	Number	Independence	Bivariate	Trivariate $P_{ij} = P_{ij}$ $ij,jj'\epsilon T'$	Exact for Equicorrelated Case $P_{k\ell} \equiv P$ all k,ℓ	Trivariate $P_{ij} = (P_{ij})^2$ $ij,jj'\epsilon T'$
.900	3	0.01777	0.01210	0.01121	0.01121	0.01183
.900	4	0.02363	0.01517	0.01339	0.01293	0.01452
.900	5	0.02944	0.01824	0.01557	0.01436	0.01732
.900	6	0.03523	0.02131	0.01775	0.01559	0.02005
.900	7	0.04098	0.02438	0.01993	0.01667	0.02284
.900	8	0.04669	0.02745	0.02211	0.01764	0.02556
.900	9	0.05237	0.03052	0.02429	0.01852	0.02835
.900	10	0.05802	0.03359	0.02648	0.01933	0.03104
.900	11	0.06364	0.03666	0.02866	0.02007	0.03383
.900	12	0.06922	0.03973	0.03084	0.02076	0.03652
.900	13	0.07476	0.04281	0.03302	0.02141	0.03932
.900	14	0.08028	0.04588	0.03520	0.02202	0.04206
.900	15	0.08576	0.04895	0.03738	0.02259	0.04474
.900	16	0.09121	0.05202	0.03956	0.02314	0.04755
.900	17	0.09662	0.05509	0.04174	0.02365	0.05024
.900	18	0.10201	0.05816	0.04392	0.02414	0.05302
.900	19	0.10736	0.06123	0.04610	0.02462	0.05571
.900	20	0.11268	0.06430	0.04828	0.02507	0.05850
.990	3	0.01777	0.00800	0.00756	0.00756	*
.990	4	0.02363	0.00902	0.00814	0.00796	*
.990	5	0.02944	0.01004	0.00872	0.00827	*
.990	6	0.03523	0.01106	0.00930	0.00852	*
.990	7	0.04098	0.01208	0.00987	0.00872	*
.990	8	0.04669	0.01310	0.01045	0.00890	*
.990	9	0.05237	0.01412	0.01103	0.00906	*
.990	10	0.05802	0.01514	0.01161	0.00920	*
.990	11	0.06364	0.01616	0.01219	0.00932	*
.990	12	0.06922	0.01718	0.01277	0.00943	*
.990	13	0.07476	0.01821	0.01335	0.00954	*
.990	14	0.08028	0.01923	0.01393	0.00964	*
.990	15	0.08576	0.02025	0.01450	0.00973	*
.990	16	0.09121	0.02127	0.01508	0.00981	*
.990	17	0.09662	0.02229	0.01566	0.00989	*
.990	18	0.10201	0.02331	0.01624	0.00996	*
.990	19	0.10736	0.02433	0.01682	0.01003	*
.990	20	0.11268	0.02535	0.01740	0.01010	*

"" means results were unobtainable due to inability to integrate accurately.

Table 2A

Bivariate Method Upper bound for z_0 so that $\Pr\{\bigcup_{i=1}^{n} X_i \notin (-z_0\sigma_i, z_0\sigma_i)\} \leq .05$

where $X \sim N(0,\Sigma)$ and there is a tree T connecting the X_i, $i=1,\ldots,n$

such that $|P_{x_i x_j}| \equiv P$ for all $e_{ij} \varepsilon T$.

n	Independence	P=.1	P=.2	P=.3	P=.4	P=.5	P=.6
4	2.494	2.493	2.491	2.486	2.478	2.467	2.449
5	2.572	2.572	2.570	2.565	2.558	2.546	2.529
6	2.635	2.635	2.633	2.628	2.621	2.610	2.593
7	2.686	---	2.685	2.681	2.674	2.663	2.646
8	2.731	---	2.730	2.726	2.719	2.709	2.692
9	2.769	---	2.769	2.765	2.759	2.748	2.732
10	2.803	---	2.803	2.800	2.793	2.783	2.767
11	2.834	---	2.834	2.831	2.825	2.815	2.799
12	2.862	---	2.862	2.859	2.853	2.843	2.828
13	2.887	---	2.887	2.884	2.876	2.869	2.854
14	2.910	---	---	2.908	2.902	2.893	2.878
15	2.932	---	---	2.929	2.924	2.915	2.900
16	2.952	---	---	2.950	2.944	2.936	2.921
17	2.970	---	---	2.968	2.963	2.955	2.940
18	2.988	---	---	2.986	2.981	2.973	2.958
19	3.004	---	---	3.003	2.998	2.990	2.976
20	3.020	---	---	3.018	3.014	3.005	2.992

"---" means value is higher than that given by conservative assumption of independence.

Table 2A Continued.

	P=.7	P=.8	P=.9	P=.95	P=.99
4	2.423	2.381	2.306	2.239	2.105
5	2.502	2.458	2.376	2.296	2.145
6	2.566	2.521	2.436	2.349	2.181
7	2.619	2.574	2.487	2.397	2.215
8	2.665	2.620	2.531	2.439	2.246
9	2.705	2.660	2.571	2.248	2.274
10	2.741	2.696	2.606	2.510	2.301
11	2.773	2.728	2.638	2.541	2.327
12	2.802	2.758	2.668	2.560	2.350
13	2.824	2.785	2.695	2.596	2.372
14	2.853	2.804	2.719	2.620	2.393
15	2.876	2.832	2.742	2.643	2.413
16	2.897	2.853	2.764	2.664	2.432
17	2.916	2.873	2.784	2.684	2.450
18	2.935	2.892	2.803	2.703	2.467
19	2.952	2.910	2.821	2.721	2.483
20	2.968	2.926	2.838	2.738	2.499

Table 2B

Bivariate Method Upper Bound for z_0 so that $\Pr\{\bigcup_{i=1}^{n} X_i \notin (-z_0\sigma_i, z_0\sigma_i)\} \leq .10$ where $X \sim N(0,\Sigma)$ and there is a tree T connecting the X_i, $i=1,\ldots,n$ such that $|P_{x_i x_j}| \equiv P$ for all $e_{ij} \varepsilon T$.

n	Independence	P=.1	P=.2	P=.3	P=.4	P=.5	P=.6
4	2.234	2.233	2.229	2.222	2.211	2.196	2.174
5	2.319	2.319	2.315	2.309	2.298	2.283	2.261
6	2.386	---	2.384	2.378	2.368	2.353	2.331
7	2.442	---	2.441	2.435	2.426	2.411	2.390
8	2.490	---	2.490	2.484	2.475	2.461	2.440
9	2.531	---	---	2.526	2.517	2.504	2.483
10	2.568	---	---	2.564	2.555	2.542	2.522
11	2.601	---	---	2.597	2.589	2.576	2.556
12	2.630	---	---	2.627	2.619	2.607	2.587
13	2.657	---	---	2.655	2.647	2.635	2.616
14	2.682	---	---	2.680	2.673	2.661	2.642
15	2.705	---	---	2.704	2.696	2.685	2.666
16	2.726	---	---	2.725	2.718	2.707	2.688
17	2.746	---	---	2.746	2.739	2.727	2.709
18	2.765	---	---	2.765	2.758	2.747	2.729
19	2.783	---	---	2.783	2.776	2.765	2.747
20	2.799	---	---	2.799	2.793	2.782	2.765

*"---" means value is higher than that give by conservative assumption of independence.

Table 2B Continued.

	P=.7	P=.8	P=.9	P=.95	P=.99
4	2.142	2.093	2.008	1.930	1.793
5	2.228	2.177	2.085	1.997	1.835
6	2.298	2.246	2.150	2.055	1.873
7	2.357	2.304	2.206	2.106	1.908
8	2.407	2.355	2.254	2.151	1.941
9	2.451	2.399	2.298	2.192	1.972
10	2.490	2.438	2.336	2.229	2.001
11	2.525	2.473	2.371	2.263	2.028
12	2.557	2.505	2.403	2.294	2.054
13	2.585	2.534	2.433	2.323	2.078
14	2.612	2.561	2.460	2.350	2.100
15	2.636	2.586	2.485	2.375	2.122
16	2.659	2.609	2.508	2.398	2.143
17	2.680	2.631	2.530	2.420	2.162
18	2.700	2.651	2.551	2.440	2.181
19	2.719	2.670	2.570	2.460	2.200
20	2.737	2.688	2.589	2.478	2.216

Table 2C

Bivariate Method Upper Bound for z_0 so that $\Pr\{\bigcup_{i=1}^{n} X_i \notin (-z_0\sigma_i, z_0\sigma_i)\} \leq .20$ where $X \sim N(0,\Sigma)$ and there is a tree T connecting the X_i, $i=1,\ldots,n$ such that $|P_{x_i x_j}| \equiv P$ for all $e_{ij} \in T$.

n	Independence	P=.1	P=.2	P=.3	P=.4	P=.5	P=.6
4	1.943	1.941	1.935	1.926	1.911	1.891	1.863
5	2.036	---	2.033	2.023	2.009	1.989	1.960
6	2.111	---	2.109	2.100	2.087	2.067	2.039
7	2.172	---	---	2.164	2.151	2.132	2.104
8	2.224	---	---	2.218	2.205	2.187	2.160
9	2.269	---	---	2.265	2.253	2.235	2.204
10	2.309	---	---	2.306	2.294	2.277	2.251
11	2.344	---	---	2.343	2.332	2.315	2.289
12	2.376	---	---	2.376	2.365	2.348	2.324
13	2.406	---	---	2.406	2.395	2.379	2.355
14	2.432	---	---	---	2.423	2.408	2.384
15	2.457	---	---	---	2.449	2.434	2.410
16	2.480	---	---	---	2.473	2.458	2.435
17	2.502	---	---	---	2.495	2.480	2.458
18	2.522	---	---	---	2.516	2.501	2.479
19	2.541	---	---	---	2.536	2.521	2.499
20	2.559	---	---	---	2.554	2.540	2.518

*"---" Means value is higher than that given by conservative assumption of independence.

Table 2C Continued.

	P=.7	P=.8	P=.9	P=.95	P=.99
4	1.823	1.765	1.668	1.581	1.433
5	1.920	1.859	1.753	1.654	1.477
6	1.999	1.936	1.825	1.718	1.518
7	2.064	2.001	1.887	1.774	1.556
8	2.120	2.058	1.941	1.825	1.592
9	2.169	2.107	1.990	1.870	1.626
10	2.213	2.150	2.033	1.912	1.657
11	2.251	2.190	2.072	1.949	1.687
12	2.286	2.225	2.108	1.984	1.715
13	2.318	2.258	2.140	2.016	1.741
14	2.347	2.287	2.171	2.046	1.767
15	2.374	2.315	2.199	2.074	1.791
16	2.399	2.340	2.225	2.100	1.813
17	2.423	2.364	2.249	2.124	1.835
18	2.444	2.387	2.272	2.147	1.856
19	2.465	2.408	2.293	2.169	1.876
20	2.484	2.427	2.314	2.189	1.895

Table 3

Selected Trivariate and Bivariate Upper Bounds for z_0 so that $\Pr\{\bigcup_{i=1}^{n} X_i \notin (-z_0\sigma_i, z_0\sigma_i)\} \leq .10$ where $X \sim N(0,\Sigma)$ and there is a tree T connecting the X_i: $i=1,\ldots,n$ such that $|P_{x_i x_j}| \equiv P$ for all $e_{ij} \varepsilon T$.

n=5

P_{ij}	Independence	Bivariate	Trivariate $(P_{ij'}=P_{ij})$ $e_{ij',jj'}\varepsilon T'$	Trivariate $(P_{ij'}=(P_{ij})^2)$ $e_{ij',jj'}\varepsilon T'$
.6	2.32	2.26	2.23	2.25
.7	2.32	2.23	2.17	2.22
.8	2.32	2.18	2.11	2.17
.9	2.32	2.09	2.02	2.07

n=10

P_{ij}	Independence	Bivariate	Trivariate $(P_{ij'}=P_{ij})$ $e_{ij,jj'}\varepsilon T'$	Trivariate $(P_{ij'}=(P_{ij})^2)$ $e_{ij,jj'}\varepsilon T'$
.6	2.57	2.52	2.49	2.50
.7	2.57	2.49	2.43	2.48
.8	2.57	2.44	2.35	2.42
.9	2.57	2.33	2.24	2.30

n=20

P_{ij}	Independence	Bivariate	Trivariate $(P_{ij'}=P_{ij})$ $e_{ij,jj'}\varepsilon T'$	Trivariate $(P_{ij'}=(P_{ij})^2)$ $e_{ij,jj'}\varepsilon T'$
.6	2.80	2.77	2.76	2.77
.7	2.80	2.74	2.73	2.74
.8	2.80	2.69	2.62	2.68
.9	2.80	2.59	2.49	2.58

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 407	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Deriving And Applying Improved Upper Bounds For Multivariate Normal Probability Outside Of N-Cubes		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Donald R. Hoover		8. CONTRACT OR GRANT NUMBER(s) N00014-86-K-0156
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 1111		12. REPORT DATE July 19, 1988
		13. NUMBER OF PAGES 81
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Bonferroni Bounds; Inclusion/Exclusion; Multivariate Normal Applications; Simultaneous Confidence Intervals

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

PLEASE SEE FOLLOWING PAGE.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

TECHNICAL REPORT NO. 407

20. ABSTRACT

Two related improved Bonferroni methods to obtain upper bounds for the probabilities of unions of events are described. One procedure (bivariate method) uses probabilities of events and pairwise intersections of events. The other superior and more complicated procedure (trivariate method) uses probabilities of pairwise and three way intersections of events. These new methods are applied to multivariate normal hypothesis testing and simultaneous unbiased confidence intervals and are shown to give results superior to those of the currently used procedure (conservative assumption of independence of events) if the number of variables is not too large and the data is highly correlated. Modifications of these new methods to non-normal probabilities and different types of probability regions are also mentioned. Keywords: Statistical [illegible]